



Faculteit Letteren & Wijsbegeerte
Departement Taalkunde

A TYPOLOGY OF NON-SELECTIVE INTERROGATIVE PRONOMINALS

(Een typologie van niet-selectieve interrogatieve pronomina)

Proefschrift voorgelegd tot het behalen van de graad
van doctor in de taal- en letterkunde aan de
Universiteit Antwerpen te verdedigen door

Dmitry IDIATOV

Promotor: Prof. Johan van der Auwera

Antwerpen, 2007

Acknowledgements

First of all, I would like to express my deep gratitude to all the people who made time to share with me their knowledge of the languages they study and/or are native speakers of. They are too numerous to be mentioned in person here but most of them are cited at various occasions in the text. I beg pardon in advance from those whose names might be missing because of my forgetfulness. The names of the people who responded to my questionnaire are summarized in Appendix A. I am also indebted to all the linguists who provided their highly valuable comments on the sections of this thesis dealing with the languages and linguistic groups they specialize in. They are named at the beginning of the respective sections.

This thesis had not been possible without the help of my supervisor Johan van der Auwera who, to begin with, believed that a Russian Africanist can also do typology in Belgium and gave me the opportunity to conduct my research in excellent working conditions at the University of Antwerp. He was also always there for me, as regards both scientific and practical issues. Furthermore, I very much appreciated the research freedom that I enjoyed all these years in my work under his supervision.

I also wish to gratefully acknowledge the financial support of the Research Council of the University of Antwerp which has been provided to me within the the GOA (Geconcerteerde Onderzoeksactie) “Mood and Modality” project.

I am truly thankful to Valentin Vydrine who introduced me into the world of linguistics and linguists, has been guiding me for many years, and opened many possibilities. He also initiated me into the joys of field work and demonstrated the (often easily overlooked) exciting sides of lexicography. I also wish to thank Thomas Bearth for making many things possible and introducing me to the Toura.

The person I am most grateful and indebted in all respects is de Makke (Mark Van de Velde). Among many other things, my gratitude goes to him for proofreading the whole text, improving my English, managing to understand what I mean even when my wording was very far from the ideals of lucidity, discussing various problematic issues, patiently waiting for the moment when the text will be ready and I will finally paint the staircase, and for making the life fun in general.

Last but not least, I wish to thank the family, friends and colleagues for making the last four and something years a pleasant time beyond the thesis.

Contents



Acknowledgements.....	i
Contents.....	iii
Abbreviations and conventions.....	xi
I. Introduction.....	1
1. Overview.....	1
2. The object of study.....	3
2.1. Pronominal.....	3
2.2. Interrogative.....	5
2.3. Selection vs. non-selection.....	7
2.4. Persons vs. things.....	11
2.5. Identity vs. classification and proper names vs. common nouns.....	16
2.6. Summary.....	17
3. Research questions.....	19
4. What qualifies as a lack of differentiation between ‘who?’ and ‘what?’ (and what does not).....	20
4.1. Simple cases.....	20
4.2. Complications.....	21
4.2.1. Syntactic function: predicate vs. non-predicate and case.....	22
4.2.2. Gender and number.....	24
4.2.3. Non-prototypical combinations of values (with respect to the choice of a non-selective interrogative pronominal).....	25
4.2.3.1. A non-prototypical combination of values vs. lack of differentiation	25
4.2.3.2. Non-human animates: a border case.....	28
4.2.3.3. Neither a lack of differentiation nor a non-prototypical combination of values.....	30
4.2.3.3.1. “No conjecture”-contexts.....	30
4.2.3.3.2. Groups of people.....	31
4.2.3.3.3. Deceptive appearances.....	32
4.2.3.3.4. Special discourse forms.....	35
5. Earlier work.....	36
6. The sample.....	38
7. Gathering the data: published sources and the questionnaire.....	40

II. Non-prototypical combinations of values	43
1. Introduction	43
2. ‘What?’-dominance in KIND-questions	51
2.1. Contexts	51
2.1.1. Biological affiliation: sex	51
2.1.2. Social affiliation	53
2.1.3. Functional affiliation	53
2.1.4. Personal affiliation	54
2.2. Constructions	56
2.3. ‘What?’-dominance in KIND-questions: summary	59
3. ‘Who?’-dominance in NAME-questions	61
3.1. Preliminaries	61
3.2. Constructions	63
3.3. Contexts	67
3.3.1. Proper names of persons	67
3.3.2. Proper names of domestic animals	74
3.3.3. Proper names of places (toponyms)	75
3.3.4. Temporal proper names	81
3.3.5. Names of “folk genera” (species)	83
3.3.6. Proper names of inanimate things	88
3.3.7. Pure autonyms	90
3.4. ‘Who?’-dominance in NAME-questions: summary	91
4. ‘Who?’-dominance in ANIMATE-questions	95
4.1. ANIMATE-questions: three strategies	95
4.1.1. ‘What?’	95
4.1.2. A special interrogative	95
4.1.3. ‘Who?’	100
4.2. Animacy: cut-off points	110
4.2.1. Cut-off points: in general	110
4.2.2. Cut-off points: in questions	111
5. Non-prototypical combinations of values: concluding remarks	117
III. Lack of differentiation	127
1. Introduction	127
2. Africa and the Middle East	129
2.1. Niger-Congo languages	129
2.1.1. Bantu languages	129
2.1.1.1. Introduction	129

2.1.1.2. Mongo (Bantu C70).....	131
2.1.1.2.1. The interrogatives <i>ná</i> and <i>é</i> : nominal uses.....	132
2.1.1.2.2. The interrogatives <i>ná</i> and <i>é</i> : other uses.....	138
2.1.1.2.3. Possible origins of the lack of differentiation between ‘who?’ and ‘what?’ in Mongo (and some other languages of zone C).....	139
2.1.2. Atlantic languages: Bijogo.....	141
2.2. “Khoisan” languages.....	143
2.2.1. Genetic classification of “Khoisan” languages.....	143
2.2.2. Eastern !Xóǝ (“Khoisan”, Tuu, Taa).....	145
2.2.3. †Hǝ (“Khoisan”, isolate or Ju).....	149
2.3. Afro-Asiatic languages.....	150
2.3.1. Introduction.....	150
2.3.2. Berber languages.....	155
2.3.2.1. Some Berber languages with ‘who?, what?’ interrogatives.....	155
2.3.2.2. Berber interrogative pronominals as constructions: simplicity of structure & multitude of forms.....	159
2.3.2.2.1. Preliminaries.....	159
2.3.2.2.2. The interrogatives <i>ma</i> , <i>may</i> , <i>mad</i> , and the like.....	161
2.3.2.2.3. The interrogatives <i>mi</i> and <i>wi</i>	164
2.3.2.2.4. The interrogatives <i>man</i> , <i>mani</i> , <i>an</i> , and the like.....	169
2.3.2.2.5. The interrogatives <i>matta</i> , <i>matak</i> , <i>ta’k(k)</i> , and the like.....	172
2.3.2.3. Berber languages: summary.....	178
2.3.3. Egyptian.....	180
2.3.3.1. Introduction.....	180
2.3.3.2. Old and Middle Egyptian.....	182
2.3.3.2.1. The interrogative <i>m(i)</i>	186
2.3.3.2.2. The interrogatives <i>pw</i> , <i>ptr</i> , <i>pty</i> , and the like.....	191
2.3.3.2.3. The interrogative <i>ĩsst/ĩšsy</i>	194
2.3.3.2.4. The interrogative <i>zy</i> and the like.....	197
2.3.3.3. Late Egyptian.....	199
2.3.3.3.1. The interrogatives <i>nym</i> and <i>m</i>	200
2.3.3.3.2. The interrogative <i>ih</i>	200
2.3.3.3.3. The interrogative <i>it(i)</i>	202
2.3.3.4. Demotic Egyptian.....	203
2.3.3.5. Coptic Egyptian.....	204
2.3.3.5.1. The interrogative <i>ou</i>	206
2.3.3.5.2. The interrogative <i>nim</i>	207
2.3.3.5.3. The interrogative <i>aš</i>	208
2.3.4. Semitic languages.....	210
2.3.4.1. Semitic interrogative pronominals: an overview.....	211
2.3.4.1.1. Forms, meanings, history.....	211

2.3.4.1.2. Deictic “reinforcement”	218
2.3.4.1.3. When ‘who?’ is <i>not</i> ‘what?’ and ‘what?’ is <i>not</i> ‘who?’	219
2.3.4.1.3.1. NAME-questions.....	220
2.3.4.1.3.2. KIND-questions.....	221
2.3.4.1.3.3. The ‘what’s up?’-construction.....	223
2.3.4.1.3.4. “Things denoting persons”	224
2.3.4.2. Semitic languages with ‘who?, what?’ interrogatives.....	228
2.3.4.2.1. Modern Arabic varieties.....	228
2.3.4.2.2. Soqotri.....	230
2.3.4.2.3. The “Canaano-Akkadian” mixed language.....	235
2.3.5. Cushitic languages.....	236
2.3.5.1. Cushitic interrogative pronominals: an overview.....	237
2.3.5.2. Saho: an East Cushitic language with a ‘who?, what?’ interrogative.....	239
3. Eurasia.....	243
3.1. Indo-European.....	243
3.1.1. Introduction.....	243
3.1.2. East Baltic languages.....	244
3.1.3. French.....	247
3.1.3.1. Modern French <i>quel?</i>	247
3.1.3.2. Middle and Classical French, non-standard French varieties and French-based creoles: the interrogative <i>qui?</i>	249
3.1.4. Old English.....	255
3.1.5. Ancient Greek.....	258
3.1.6. Latin.....	261
3.2. North Caucasian languages.....	263
3.2.1. East Caucasian.....	263
3.2.2. West Caucasian.....	271
3.2.2.1. Tapanta Abaza.....	271
3.2.2.2. (Standard) Abkhaz.....	277
3.3. Dravidian languages.....	278
3.3.1. Dravidian gender systems.....	279
3.3.2. Dravidian interrogative pronominal systems.....	281
3.3.3. Dravidian languages with ‘who?, what?’ interrogatives.....	286
3.3.4. On the Proto-Dravidian interrogative pronominal system.....	291
3.4. Kusunda.....	293
3.5. Tungusic languages: Evenki.....	298
3.5.1. The interrogative <i>aŋi:</i>	299
3.5.2. The interrogatives <i>e:-</i> and <i>e:-kun</i>	303
4. Southeast Asia and Oceania.....	309
4.1. Austro-Asiatic languages.....	309

4.1.1. Introduction.....	309
4.1.2. Khasian languages.....	313
4.1.2.1. Khasi gender systems.....	313
4.1.2.2. Khasi interrogative pronominal systems.....	317
4.1.2.2.1. Standard Khasi.....	318
4.1.2.2.1.1. The interrogatives <i>no</i> and <i>ey</i>	318
4.1.2.2.1.1.1. Nagaraja (1985) vs. Rabel (1961)	319
4.1.2.2.1.1.2. Gender-number markers and the distinction human vs. non-human.....	320
4.1.2.2.1.1.3. “Prepositions” and the distinction human vs. non-human.....	323
4.1.2.2.1.1.4. The interrogatives <i>no</i> and <i>ey</i> and the distinction human vs. non-human: summary.....	327
4.1.2.2.1.2. The interrogative <i>ayu</i>	327
4.1.2.2.2. Major non-standard Khasi varieties: War, Pnar and Lyngngam....	330
4.1.2.2.2.1. War.....	330
4.1.2.2.2.2. Pnar.....	332
4.1.2.2.2.3. Lyngngam.....	333
4.1.2.3. On the etymology of the interrogatives <i>no</i> and <i>ey</i>	335
4.1.3. Palaungic languages.....	336
4.1.4. Bahnaric languages.....	340
4.1.4.1. Sapuan and Cheng.....	342
4.1.4.2. Stieng.....	346
4.1.5. Aslian languages.....	348
4.1.6. Munda languages.....	351
4.2. Austronesian languages.....	357
4.2.1. Introduction.....	357
4.2.2. Northern Luzon languages.....	360
4.2.2.1. South-Central Cordilleran languages.....	361
4.2.2.1.1. The ‘name’-based interrogatives.....	363
4.2.2.1.2. The * <i>sɿ</i> -initial interrogatives.....	365
4.2.2.1.3. Central Cordilleran languages.....	372
4.2.2.1.3.1. Kalinga languages.....	372
4.2.2.1.3.1.1. Upper Tanudan Kalinga.....	372
4.2.2.1.3.1.2. Lower Tanudan Kalinga.....	372
4.2.2.1.3.1.3. Limos Kalinga.....	373
4.2.2.1.3.1.4. Guinaang Kalinga.....	374
4.2.2.1.3.2. Nuclear Cordilleran languages.....	375
4.2.2.1.3.2.1. Lagawe and Bayninan Ifugao.....	375
4.2.2.1.3.2.2. Northern Kankanay.....	376
4.2.2.1.3.2.3. Balangao.....	377
4.2.2.1.4. Southern Cordilleran languages.....	378

4.2.2.2. Northern Cordilleran languages.....	380
4.2.3. Lampungic languages.....	383
4.2.4. Choiseul languages.....	389
4.2.5. South Huon Gulf languages.....	390
4.2.6. Tsouic languages.....	393
4.2.6.1. Kanakanabu.....	395
4.2.6.1.1. The interrogatives <i>namanaini</i> and <i>nein</i>	396
4.2.6.1.2. The interrogative <i>nian</i>	399
4.2.6.2. Saaroa.....	400
5. New Guinea: Kiwaian languages.....	403
6. Australia.....	406
6.1. Introduction.....	406
6.2. Australian languages with ‘who?, what?’ interrogatives: distribution...	409
6.3. Australian interrogative pronominals: some common functional and formal patterns.....	415
6.3.1. The interrogative/indefinite relation.....	415
6.3.2. The link between ‘where?’, ‘which [N]?', ‘which one?', ‘who?’ and ‘what?’.....	418
6.3.3. Gender marking on interrogatives.....	424
6.3.4. Shortening and augmentation.....	427
6.4. A closer look at the Australian languages with ‘who?, what?’ interrogatives.....	432
6.4.1. Pama-Nyungan languages.....	432
6.4.1.1. The interrogatives <i>ngana</i> > <i>ngana-nha</i> > <i>nhaa</i> and the like.....	432
6.4.1.1.1. Kurna and Parnkalla.....	433
6.4.1.1.2. Wirangu.....	436
6.4.1.1.3. Mirning, Ngadjunmaya and Karlamay.....	439
6.4.1.1.4. Pintupi.....	440
6.4.1.1.5. Warumungu.....	443
6.4.1.1.6. Yulparija, Walmajarri and Warlmanpa.....	446
6.4.1.1.7. Mangala.....	447
6.4.1.2. Some possible leftovers.....	448
6.4.2. Non-Pama-Nyungan languages.....	449
6.4.2.1. Form of the gender-number markers: Anindilyakwa.....	450
6.4.2.2. Gender-number semantics.....	450
6.4.2.2.1. Burarra and Worrorra.....	451
6.4.2.2.2. Wambaya.....	452
6.4.2.3. Jingulu.....	453
6.4.2.4. Nyulnyulan languages.....	456
6.4.2.5. Tangkic languages.....	461
6.4.2.6. Garrwan languages.....	466

6.4.2.7. Rembarrnga.....	471
6.5. Concluding remarks.....	473
6.5.1. On Dixon's (2002) generalizations.....	474
6.5.2. The lack of differentiation between 'who?' and 'what?': Pama-Nyungan vs. non-Pama-Nyungan.....	475
7. North America.....	479
7.1. Introduction.....	479
7.2. Nishnaabemwin (Ojibwa).....	480
7.3. Timucua.....	485
7.4. Timbisha.....	487
7.5. Mayan languages.....	492
7.5.1. Tz'utujil.....	492
7.5.2. Kaqchikel.....	496
7.5.3. Mam.....	499
7.5.4. Uspanteka.....	502
7.5.5. Tzotzil.....	503
7.5.6. Mayan languages: concluding remarks.....	504
7.6. Garifuna and Island Carib.....	511
8. South America.....	515
8.1. Introduction.....	515
8.2. Arawak(an) languages.....	517
8.2.1. Complex constructions and their truncation.....	523
8.2.2. 'Where?', 'which one?', 'who?', and 'what?'.....	529
8.2.3. Gender.....	529
8.3. Chibchan languages: Arhuaco/Ika.....	534
8.4. Makú languages: Dâw.....	535
8.5. Urarina.....	536
8.6. Arauan languages: Paumarí.....	537
8.7. Chapacura-Wanham languages: Itene/More.....	539
8.8. Tacanan languages.....	543
8.9. Tupí languages.....	545
8.10. Macro-Jê languages.....	548
8.10.1. Xerente.....	549
8.10.2. Krenák.....	552
8.10.3. Maxakalí.....	554
8.10.4. Fulniô/Yatê.....	556
8.11. Guaicuruan languages.....	556
8.12. Matacoan languages.....	559
9. Lack of differentiation: concluding remarks.....	562
9.1. Main results.....	562

9.2. Lack of differentiation between ‘who?’ and ‘what?’: how widespread it is and why.....	563
9.3. Lack of differentiation between ‘who?’ and ‘what?’: areal and genetic distribution.....	566
9.4. Lack of differentiation between ‘who?’ and ‘what?’: the many ways to get there.....	574
Appendix A: The questionnaire respondents.....	577
Appendix B.1: The questionnaire: the “regular” version.....	584
Appendix B.2: The questionnaire: the “light” version.....	587
Appendix C: The sample.....	588
Appendix D: The reduced sample.....	602
Appendix E: Dravidian interrogative pronominals.....	609
References.....	616
Nederlandse samenvatting.....	651

Abbreviations and conventions

I use the glossing conventions set forth in the Leipzig Glossing Rules (<http://www.eva.mpg.de/lingua/files/morpheme.html>), with a few additions and modifications. Examples, tables, figures, maps and footnotes are numbered continuously only in Chapter I, *Introduction*. In Chapter II, *Non-prototypical combinations of values*, and Chapter III, *Lack of differentiation*, the numbering is separate for each first-level section, such as Section II.1, Section II.2, Section II.3, etc. The Roman number refers to the chapter and the following Arabic number refers to the section within this chapter. References to examples, tables, figures, maps and footnotes within the same chapter (for Chapter I) or the same first-level section (for Chapters II and III) consist of Arabic numbers only, e.g. *see example (23)* or *see Map 2*. A reference to an example in another chapter or (not for Chapter I) in another first-level section within the same chapter begins with the number of the section (including the chapter number). For instance, *see example (II.3.1.3.1:23)* refers to example (23) in Section II.3.1.3.1. Similarly, *see Map III.6.2:2* refers to Map 2 in Section III.6.2.

...	underlying/morphonological/basic form	AG	agreement pattern (when the name of an agreement pattern is not a number, it is enclosed in <...>, e.g. AG<IJO> stands for <i>agreement pattern IJO</i>)
...A/D/E	“person inflection classes” A, D, E (in Urarina)		
A	agent-like argument of a canonical transitive verb	ALL	allative
ABL	ablative	AN	animate
ABS	absolutive	AOR	aorist
ACC	accusative	APPL	applicative
ACT	active	APOSS	alienable possession
ADJ	adjective	ART	article
ADJZ	adjectivizer	AS	“annexed state” (in Berber languages)
ADV	adverb(ial)		
AFF	affix: (i) affixes left unglossed in the sources, (ii) affixes glossed in the sources but left without any explanation of the meaning of the gloss, (iii) affixes, the exact meaning of which is irrelevant for the discussion	ASS	associative
		AUG	(in Bantu) augment, (elsewhere) augmented
		AUX	auxiliary
		BEN	benefactive
AFOC	“agent focus” (in Austronesian languages)	BFOC	“benefactive focus” (in Austronesian languages)
		CAUS	causative
		CLF	classifier
		CMPR	comparative

COM	comitative	IPRF	imperfect
COMP	complementizer	IPW	interrogative pro-word
COMPL	completive	IRR	irrealis
CON	connective	LFOC	“locative focus” (in Austrone- sian languages)
COP	copula	LIG	“ligature” (in Austronesian languages)
CS	“construct state” (in Semitic languages)	LOC	locative
DAT	dative	M	masculine
DEF	definite	MIN	minimal
DEM	demonstrative	MPO	“multipurpose oblique” (in “Khoisan” languages)
DEP	dependent (in Egyptian, “de- pendent pronouns”)	N	neuter
DIM	diminutive	NAFOC	“non-agent focus” (in Aus- tronesian languages)
DU	dual	NARR	narrative
DUB	dubitative	NEG	negation
DUR	durative	NEU	“neutral case” (in Mayrinax Atayal)
EMPH	emphatic	NMLZ	nominalizer
EP	epenthetic	NOM	nominative
ERG	ergative	NON<...>	non-<...>
EXCL	exclusive	NS	“nominal specifier” (in Aus- tronesian languages)
F	feminine	OBJ	object (usually, the patient- like argument of a canonical transitive verb)
FOC	focus	OBL	oblique
FS	“free state” (in Berber lan- guages)	OBV	obviative
FUT	future	OPT	optative
G	gender	P	predication
GEN	genitive	PASS	passive
HAB	habitual	PFOC	“patient focus” (in Austrone- sian languages)
HON	honorific	PFV	perfective
HUM	human	PL	plural
IMP	imperative	PLH	placeholder, filler (e.g., ‘whatchamacallit’, ‘whatsi(t)sname’, ‘so-and-so’, etc.)
INAN	inanimate	POSS	possessive
INCL	inclusive		
INCOMPL	incompletive		
IND	indicative		
INDEP	independent (in Egyptian, “in- dependent pronouns”)		
INDF	indefinite		
INF	infinitive		
INS	instrumental		
IPFV	imperfective		

PRED	predicate	S	single argument of a canonical intransitive verb
PRF	perfect		
PRF ₂	“the second perfect” (in Coptic)	s1	“Set I” interrogative pronominal (in Dravidian languages, cf. Section III.3.3.2)
PRN	pronoun		
PROG	progressive	s2	“Set II” interrogative pronominal (in Dravidian languages, cf. Section III.3.3.2)
PROP	proper name		
PROX	proximate		
PRS	present	SBJ	subject
PQ	polar question	SBJV	subjunctive
PST	past	SG	singular
PTCL	particle	SMLF	semelfactive
PTCP	participle	SUB	subordinate/subordinator
Q	(constituent or polar) question marker	SUF	“suffixed pronouns” (in Egyptian)
REAL	realis	TOP	topic
RED	reduplication	VEG	vegetable
REF	referential	VEN	ventive (“centripetal”)
REFL	reflexive	<i>vi</i>	intransitive verb
REL	relative	VOC	vocative
RES	resultative	<i>vt</i>	transitive verb
RN	“relational noun” (in Mayan languages)		

KIND-questions, NAME-questions, ANIMATE-questions: cf. Section II.1

I Introduction

1 Overview

This study in lexical typology investigates non-selective interrogatives pronominals equivalent to English *who?* (1) and *what?* (2).

English

- (1) a. *Who is that man over there?*
b. *Who gave you this?*
c. *Who did he kill?*
- (2) a. *What is this thing you have in your hand?*
b. *What fell out of his bag?*
c. *What can you make out of clay?*

In particular, I explore the diversity of natural (spoken) languages in the domain of the non-selective interrogatives pronominals ‘who?’ and ‘what?’ as regards (i) the universality of their formal differentiation and (ii) the patterns of their functional differentiation in languages where they are distinguished formally. This exploration is undertaken from a functional-typological perspective (as understood by Croft (2003), for instance).

The discussion is organized as follows. In the remaining introductory sections I will first delimit the object of this study, viz. non-selective interrogative pronominals meaning ‘who?’ and ‘what?’ (Section I.2). Then, I will present the research questions (Section I.3) and discuss what qualifies as a lack of differentiation between ‘who?’ and ‘what?’ and what does not (Section I.4). In Section I.5, I will provide an overview of earlier work on these issues. The general introduction will be concluded by a discussion of the sample (Section I.6) and the way my data have been collected (Section I.7).

The body of this study consists of two parts. The first part is dedicated to a

cross-linguistic investigation of patterns of functional differentiation between ‘who?’ and ‘what?’ in terms of non-prototypical combinations of values and ‘who?’ or ‘what?’-dominance, as defined in Section I.4.2.3 and Section II.1. After a few introductory remarks in Section II.1, I first discuss in Section II.2 the use of ‘what?’ in questions about the classification of a person. Then, in Section II.3 I discuss the use of ‘who?’ in questions about proper names of things and in Section II.4 the use of ‘who?’ in questions about animate things. In Section II.5, a conclusion is provided.

In the second part, I discuss various languages that appear to allow for a lack of differentiation between ‘who?’ and ‘what?’. The languages are grouped geographically in seven areas: (i) Africa and the Middle East (Section III.2), (ii) Eurasia (Section III.3), (iii) Southeast Asia and Oceania (Section III.4), (iv) New Guinea (Section III.5), (v) Australia (Section III.6), (vi) North America (down to Panama in the south and excluding the islands of the Caribbean; Section III.7), (vii) South America (Section III.8). Within these areas the languages are organized genetically. For each language I first try to determine whether we can truly speak about a lack of differentiation between ‘who?’ and ‘what?’. If the answer is positive, I attempt (within the limits of the data and time available) to establish the origins of the lack of differentiation between ‘who?’ and ‘what?’ in each particular case.

Before proceeding to the discussion, I would like to make two practical remarks. The first remark concerns my use of the names of languages and the labels for their genetic affiliations. Throughout the present study, I normally recur to the names and classifications provided by the latest edition of the *Ethnologue* (Gordon 2005). Despite all the imperfections of the *Ethnologue*, it remains for the time being the best available classification covering the overwhelming majority of the world’s languages. Note, however, that I do provide alternative classifications and names for the languages mentioned at the places where this may have some relevance for the discussion or when the nomenclature normally used in a given linguistic tradition differs considerably from the one adopted in the *Ethnologue*.

The second remark concerns my writing style. I am not a native speaker of English and although I had a lucky opportunity to have my text proofread by two Anglicists, I am aware that there is still room for improvement, especially as regards style and structure. As aptly pointed out by Nichols (1988:400), “relations among reader, writer, text, and knowledge differ in Russian and English scientific prose [...] the Russian text is not so much a communicative contract between writer and reader as it is a gnomic statement of available knowledge”. The form of my prose slowly shifted “westward” as the work progressed.

2 The object of study

In Section I.1, the object of this study has been preliminarily characterized as interrogative pronominals similar to English *who?* and *what?*. In what follows, I will elaborate on this characterization by discussing the terms and parameters *pronominal* (Section I.2.1), *interrogative* (Section I.2.2), *selection* vs. *non-selection* (Section I.2.3), *person* vs. *thing* (Section I.2.4), and *identification* vs. *classification* and *proper name* vs. *common noun* (Section I.2.5). In Section I.2.6, I will organize the relevant parameters into a single conceptual space, which will allow me to define the prototypical functions of the interrogatives ‘who?’ and ‘what?’ for purposes of cross-linguistic comparison.

2.1 Pronominal

I use the term *pronominal* to refer to a morphologically non-bound¹ substitute of a nominal, which itself does not need to belong to the lexico-syntactic class of nominals.

The term *substitute* implies that there is an equivalence relation between proform X and nominal Y that X substitutes. In the context of a constituent question, this means that the interrogative pronominal X is used to inquire about a certain referent, which in the answer to the question will be expressed by the nominal Y. In the terminology of van den Eynde & Mertens (2003:70), interrogative pronominals could be described as suspensive pronominals (“pronoms suspensifs”), since their referential specification is suspended (“la spécification référentielle est suspendue”).² The definition of an interrogative pronominal as a substitute of a nominal excludes attributive interrogatives, such as English *which [N]?*, *what [N]?* or Ewe *[N] ka* (cf. below) from the scope of this study.³ Interrogative pronominals can be non-predicative or predicative, though, such as *who?* in *Who is knocking on the door?* and in *Who is it?* respectively.

The restriction of interrogative pronominals to morphologically non-bound

¹ This includes morphologically strongly and weakly autonomous word forms and clitics and excludes “formants” and affixes (cf. Plungian 2000:32-34, who to a large extent develops on Mel’čuk 1993-2000).

² Suspensive pronominals are not necessarily interrogative. Relative pronominals may be suspensive as well (van den Eynde & Mertens 2003:70). This may account for the fact that many languages use the same forms as both interrogative and relative pronominals.

³ Some of these forms, viz. English *which?* and *what?* and Ewe *ka*, may also be used as non-selective interrogative pronominals. For instance, Ewe *ka* can also be used as ‘who?’ (cf. Westermann 1961:32).

elements is first of all intended to exclude bound roots, such as Ngandi *-nyja* (Gunwinyguan; Australia; Capell 1942:385-386; Heath 1978), which must be marked for gender, or the Tapanta Abaza interrogative root *-a*, as in (3).

Tapanta Abaza (Northwest Caucasian; Russia; Genko 1955:106)

- (3) a. *wə-z-zə-pšʃa-wa-j-a?*
 2SG.M[S]-PTCP.OBL-APPL-look-PTCP.IPFV-3SG.NON<HUM>.S-IPW
 ‘What are you looking for? (lit.: ‘The one that you are looking for, it is who/what?’)’
- b. *wə-z-zə-pšʃa-wa-d-a?*
 2SG.M[S]-PTCP.OBL-APPL-look-PTCP.IPFV-3SG.HUM.S-IPW
 ‘Who are you looking for? (lit.: ‘The one that you are looking for, s/he is who/what?’)’

This restriction is desirable because bound roots are substantially different from non-bound word forms. Thus, they are normally much less psychologically salient for an average speaker in that their meanings are less readily accessible (or identifiable) to him or her than those of non-bound word forms.⁴ In other words, they are much less part of the lexicon of a given language than non-bound word forms. What is more, at least in the domain of interrogative pronominals, bound roots tend to show fewer semantic distinctions than non-bound word forms, which makes them somewhat less interesting as an object of study. In determining whether a given form is bound or not, I normally follow the analysis adopted in the sources.

The hedge in the definition of pronominals specifying that a pronominal itself does not need to belong to the lexico-syntactic class of nouns is necessary to account for cases when questions about the identity of persons and things are expressed with conventionalized noun phrases or clausal constructions not based on nominal interrogative pronominals. The use of conventionalized noun phrases as interrogative pronominals can be exemplified with Sango *zò wà* ‘who?, which/what person? (*zà* ‘person’, *wà* ‘which [N]?, what [N]?’)’ (Creole; Central African Republic; Thornell 1997:70) or Ewe *àmé ka* ‘who?, which/what person? and *nú kà* ‘what?, which/what thing?’ (*àmé* ‘person’, *nú* ‘thing’ and *ka* ‘which [N]?, what [N]?’; Niger-Congo, Kwa; Ghana; Pasch 1995:79). The use of clausal constructions as interrogative pronominals can be illustrated with the Apurucayali Asheninca interrogative *oitarika* (4), which is a form of the verb *-i* ‘to name, to call’ literally meaning something like ‘it calls itself?’ or ‘it is called?’ (cf. Section III.8.2.1), or with the Tapanta Abaza form *dzačʷəja* ‘who? (SG)’ (5), which can be literally translated as something like ‘the one to which she/he relates as its, it is who?’ or ‘to what kind does she/he belong?’ (cf. Section

⁴ Admittedly, much depends here on the morphological type of the language.

constituent question and of its possible answer, and is formally expressed by an interrogative proform. Furthermore, to count as an interrogative pronominal in a given language, the form must have the function of a request for an instantiation of the nominal variable x as one of its conventionalized functions. Indeed, a phrase like (7) may be used to provoke the same kind of response (9) as a conventionalized constituent interrogative construction in (8), which does not mean, however, that *someone* in (7) is an interrogative pronominal in English.

English

- (7) [A:] *By the way, yesterday, I saw you in the shop talking to someone I don't know.*
- (8) [A:] *By the way, who was it I saw you in the shop talking to yesterday?*
- (9) [B:] *Oh, it was John from my French class. He is very talkative.*

Somewhat trickier are the so-called placeholders or fillers, such as English *whatchamacallit*, *whatsit*, *whatsi(t)sname*, etc. In some respect, they are very close to real interrogatives. However, the important difference is that placeholders do not constitute a conventionalized way of requesting for information. Rather, they are linguistic devices available to the speaker “for dealing with word-formulation trouble” (Hayashi & Yoon 2006:485). For similar reasons, I do not consider as interrogative pronominals relative pronominals or pronominals used to form indirect questions.

It may happen, of course, that a language may conventionalize one and the same form not only as an interrogative pronominal but also, for instance, as an indefinite pronominal, as is typical in Australian languages (cf. Dixon 2002:328; Mushin 1995), as illustrated in (10) with Yidiny *wanyju* ‘who?, someone’.

-
- (ii) *Who has bitten into the apple?*

In (ii) there is a clear existential presupposition that there is somebody who has bitten into the apple, whereas (i) only supposes that there possibly is somebody who can solve the problem and whose identity the hearer is invited to disclose provided, of course, the latter knows such a person. If even such a possibility were not supposed, that is if the presupposition had an *It is known that it is not possible that* HAPPEN/EXIST (... x ...) structure, we would rather deal with a rhetorical question as in (iii). However, rhetorical questions are rather a matter of pragmatics than semantics.

- (iii) *Who can solve this problem?! [It is simply impossible!]*

Yidiny (Pama-Nyungan; Australia; Dixon 1977:182, 2002:328)

- (10) *wanyju walba yanggi:ny*
 who/someone.ERG rock.ABS split-PST
 ‘Someone must have cut the rock’ or ‘Who cut the rock?’, or ‘Someone must have cut the rock – who did it?’

At least in certain contexts, the same is also possible in various Indo-European languages, as illustrated by Russian *čto*, which in (11a) is used as ‘anything’, but is otherwise also a regular interrogative pronominal ‘what?’ (11b).

Russian

- (11) a. *esli tebe čto nado, daj mne znat’*
 if to.you anything be.necessary let me know
 ‘If you need anything, let me know.’
 b. *čto tebe nado?*
 what to.you is.necessary
 ‘What do you need?’

Whether such cases are to be considered as instances of polysemy or homonymy is of little relevance for the present study. However, only examples like (10) or (11b) but not (11a) will be taken into consideration.

2.3 Selection vs. non-selection

Interrogative pronominals can be used in selective contexts, when the speaker perceives the choice as being restricted to a closed set of similar alternatives (12), or in non-selective contexts, where the speaker perceives the choice as free (13) or (14).

English

- (12) **Which** (*one*) should I take? *This, that, or maybe that?*
 (13) **What** have you liked most about this book?
 (14) **Who** do you see there?

Selective interrogative pronominals typically expect a deictic nominal as an answer, such as *this (one)* or *that (one)*, rather than a descriptive or proprial nominal. In other words, selective interrogative pronominals may be characterized as something like interrogative prodeictics (or prodemonstratives). With non-selective interrogatives, the situation is the other way around. Indeed, the answer *this (one)* or *that (one)* in a non-selective context, such as (13) or (14), would be completely infelicitous.

Although many contexts may be clearly either selective or non-selective, the distinction between selective and non-selective contexts is not always easy to

make. The reason is that the decision on whether the choice is restricted or free is often a matter of the speaker's conceptualization of reality rather than of reality itself. For instance, one would normally ask *What day (of the week) is it today?* rather than *Which day (of the week) is it today?*, even though it is clear that the speaker expects no more than seven possible answers, viz. *Monday, Tuesday, Wednesday, Thursday, Friday, Saturday or Sunday*.⁶

One of the apparently selective contexts, where many languages will still prefer or at least allow a non-selective interrogative, is when the choice is asked to be made between entities of different kinds. For instance, compare the translations in different languages of verse 23:17 from The Gospel of Matthew in (15-18), where the choice is to be made between 'gold' and 'temple'.

English

- (15) a. *Ye fools and blind: for **whether** is greater, the gold, or the altar that sanctifieth the gold?* (King James Version)
 b. *Fools and blind! For **which** is greater, the gold or the temple that sanctifies the gold?* (New King James Version)
 c. *Ye fools and blind, for **what** is greater, the gold, or the temple that halloweth the gold?* (Wycliffe New Testament)

Dutch

- (16) *Gij dwazen en blinden, want **wat** is meerder, het goud, of de tempel, die **what** is greater het goud heiligt?* (Staten Vertaling)

German

- (17) *Ihr Narren und Blinden! **Was** ist größer: das Gold oder der Tempel, der **what** is greater das Gold heiligt?* (Luther 1912; also was 'what?' in the following versions: Elberfelder 1871, 1905; Luther 1545; Schlachter 1951)

French

- (18) a. *Insensés et aveugles! **lequel** est le plus grand, l'or, ou le temple qui **which.one** is the most great sanctifie l'or?* (Louis Second 1910)

⁶ The distribution between the attributive non-selective *what [N]?* and selective *which [N]?* in English is a complex issue, which is further complicated by the fact *what [N]?* can also mean 'what kind of [N]?'. English is clearly not exceptional in this respect. Consider, for instance, Choi's (2005) corpus-based study of the Korean attributive interrogatives *mwusun* 'what [N]?' and *enu* 'which [N]?'.

b. *Insensés et aveugles que vous êtes! **Qu'est-ce qui** est plus important:
what is more important
 l'or ou le Temple qui rend cet or sacré? (La Bible du Semeur)*

It is clearly not a coincidence either that it is hardly possible to answer questions like this with a deictic nominal, such as *this (one)* or *that (one)*, even when the options are visibly present both to the speaker and the interlocutor, (19) vs. (20).

English

- (19) [An apple and an orange are lying on the table. A:] *Which one/what would you like, an apple or an orange?*
 a. [B:] *An apple, please.*
 b. [B:] [?] *This one, please.*
- (20) [Two apples are lying on the table. A:] *Which one would you like?*
 [B:] *This one, please.*

Answer (19b) would be more acceptable, though, if the alternatives, *an apple or an orange*, were not expressed in the question.

The differentiation between selective and non-selective interrogative pronominals is further complicated by the fact that on the one hand, interrogative pronominals typically used as non-selective may sometimes be used selectively, whereas on the other hand, selective interrogative pronominals may sometimes be used non-selectively. Thus, it appears that in the languages of the world, human interrogatives, such as English *who?*, are regularly used in selective questions about humans instead of the dedicated selective pronominals, such as English *which one?*. On the contrary, the use of non-human interrogatives, such as English *what?*, in selective questions about things appears to be much less common. This can be illustrated with the English examples in (21).

English

- (21) a. *All these women here... and who/which is Mary?*
b. *All these cars here... and which[?] what is yours?*

Similarly, when selective interrogative pronominals are used non-selectively, it appears to be cross-linguistically much more common for an otherwise selective interrogative pronominal to be used non-selectively in questions about humans as ‘who?’, rather than in questions about things as ‘what?’. For instance, in Latvian, which normally uses a general non-selective interrogative pronominal *kas* both in the meaning ‘who?’ and ‘what?’, the selective interrogative *kurš* ‘which (one)? (M.SG)’ may also be used non-selectively, but only as the human interrogative ‘who?’ (cf. Holst 2001:131; Nau 1999:145-147). In Machiguenga (Arawakan; Peru; Betty Snell, p.c.), the animate interrogative *tyani* ‘where is he/she/it (AN)?’

which one (AN)?’ (lit.: be.where?-AN) is also used as the non-selective human interrogative ‘who?’, instead of the dedicated ‘who?’ interrogative *tsini*, which nowadays is used only infrequently. At the same time, *tyati*, the inanimate counterpart of *tyani*, appears to be used only selectively as ‘which one (INAN)?’.⁷

This preferential link between ‘which one?’ and ‘who?’ rather than ‘what?’ in the languages of the world is just one more example of the strong correlation between humanness and higher degree of individuation and referentiality, as is also reflected in the so-called Extended Animacy Hierarchy (cf., e.g., Croft 2003:131-132). The same correlation can also be found in the typical answer patterns related to ‘who?’ and ‘what?’, cf. Section I.2.5.

The present study is confined to non-selectively used interrogative pronominals. This restriction is justified by the fact that in many respects selective interrogative pronominals are substantially different. Thus, they are specialized for use in different contexts, which also manifests itself in the fact that basically, they expect a deictic nominal as an answer, rather than a descriptive or proprial one. Accordingly, as compared to non-selective interrogative pronominals, they tend to show different semantic and morphosyntactic distinctions, such as number, gender (rather than human vs. non-human distinction)⁸ or number of alternatives involved in the selection (usually, two alternatives only vs. two or more alternatives).⁹

I refer to non-selective interrogative pronominals as ‘who?’ for questions about humans and as ‘what?’ for questions about non-humans respectively. It should be noted, though, that the use of the glosses ‘who?’ and ‘what?’ does not imply that the interrogative at issue cannot be used selectively as well (cf.

⁷ The locative interrogative ‘where?’ unmarked for animacy is *tya-ra* (Snell 1998:70-71). The final *-ra* here seems to be the locative nominalizer, as in *no-mag-i-ra* (1SG-sleep-REAL-LOC.NMLZ) ‘where I sleep, my bed’ or *i-nori-a-ra* ‘where he is lying down’ (3M-lie.down-REAL.REFL-LOC.NMLZ) (Snell 1998:68-69).

⁸ Recall that as suggested earlier in this section, selective interrogative pronominals are probably better characterized as interrogative pro-deictic nominals. Accordingly, it is no surprise that in the languages of the world the explicit differentiation between persons and things is equally uncommon in both demonstratives (cf. Diessel 2003:641-643) and selective interrogative pronominals, alias interrogative pro-deictic nominals. This contrasts radically with the fact that in non-selective interrogative pronominals an explicit differentiation between persons and things is extremely widespread cross-linguistically.

⁹ For instance, Lithuanian (Indo-European, Baltic; Ambrazas 1997:198-199, 209-212) has the following selective interrogative pronominals: *katràs* (M.SG.NOM)/ *katrà* (F.SG.NOM)/ *katriẽ* (M.PL.NOM)/ *katros* (F.PL.NOM) ‘which one(s) of two?’ vs. *kuris* (M.SG.NOM)/ *kuri* (F.SG.NOM)/ *kuriẽ* (M.PL.NOM)/ *kuriõs* (F.PL.NOM) ‘which one(s)? (“from a group of two or any limited number”)’.

above). Selective interrogative pronominals are normally referred to as ‘which one?’, with an optional specification (person or thing) between brackets. The gloss ‘which (one)?’ implies that the selective interrogative can be both a pronominal and an attribute. Attributive interrogatives are referred to as ‘which [N]?’ if selective, and ‘what [N]?’ if non-selective. The latter should not be mixed up with qualifying ‘what kind of [N]?’. It should be mentioned, however, that in many cases, it happened to be hardly possible to determine the exact meaning of the attributive interrogative, which is due to both the complexity of the issue itself and the fact that it is rarely treated in any detail in grammatical descriptions. Thus, I often had to take over the gloss ‘which?’ from the sources, which does not specify whether the interrogative is used only attributively or can be a pronominal as well, whether it is only selective or can be non-selective as well, and whether it is applicable to both things and persons or only things (or only persons).¹⁰

2.4 Persons vs. things

As pointed out by Wierzbicka (1996:38), “the distinction between [...] ‘person’ and ‘thing’ provides the most fundamental form of human categorization” of entities. It may be more correct, however, to say that the distinction is between PERSON and NON-PERSON, with THING being just the prototypical instance of the latter category. Other cut-off points, e.g. based on animacy, are secondary and “[w]hat has been referred to in the literature as the animacy hierarchy is [...] essentially a reflection of different ways of realizing grammatically a fuzzy dichotomy, at the base of which is the distinction between persons, that is, essentially human beings perceived as agents, and the rest of the universe” (Dahl & Fraurud 1996:62). The division between the concepts of PERSON and NON-PERSON may be fuzzy “because we have the possibility of sometimes treating inanimate entities as persons and, perhaps less often, human beings as non-persons, in one sense or another” (Dahl & Fraurud 1996:62). Thus, on the one hand, languages may extend the category PERSON to nominals denoting dead people, groups of people, organizations, animals, mythological beings, stars, people-like or exceptional things, etc. On the other hand, languages may extend the category of NON-PERSON to some nominals denoting humans, such as nominals denoting children or some exceptional, abnormal humans. Still, “despite all the differences in cultural context and cultural interpretation [...]”,

¹⁰ I have the impression, though, that if the distinction between persons and things is applicable to an interrogative glossed as ‘which?’ (or ‘which one?’), it would almost always be explicitly mentioned in the source. That is, it seems safe to assume that by default, an interrogative glossed as ‘which?’ (or ‘which one?’) may be used for both things and persons.

the concept PERSON has “a stable irreducible core across all languages and cultures” (Wierzbicka 1996:39).

The fundamental nature of the distinction between PERSON and THING also raises the question of whether the two concepts can be defined in a more precise way or whether they are what Wierzbicka calls semantic primitives (or primes), i.e. semantic elementary particles “in terms of which all complex meanings can be coherently represented”, but which “cannot be defined themselves” (Wierzbicka 1996:10). At first sight, PERSON (and THING) appears to be a good candidate for the status of semantic primitive. However, Wierzbicka (1996:41) argues that “the notion of an individual human being does not need to be regarded as primitive”, rather as “cross-linguistic evidence suggests” it is the concept of PEOPLE, “a social, rather than biological, category [...] which is indeed universal”. This would imply that the concept PERSON should be defined via the concept PEOPLE, viz. probably something like ‘one of people’ or ‘one of the kind of people’ (all the elements in these definitions are also putative semantic primitives).¹¹ At the same time, Wierzbicka (1996) also suggests the following two semantic primitives, SOMEONE/WHO and SOMETHING/WHAT/THING.¹² The slash here divides the so-called “allolexes” of the semantic primitive, i.e. “by analogy with ‘allomorphs’ and ‘allophones’ [...] different exponents of the same primitive” (Wierzbicka 1996:26).

What strikes the eye here is the obvious asymmetry between these two semantic primes, viz. SOMEONE/WHO rather than SOMEONE/WHO/PERSON, on the one hand, and SOMETHING/WHAT/THING, on the other. In fact, Wierzbicka herself at several places refers to ‘person’ as an allolex of SOMEONE or uses ‘person’ as a clarification of SOMEONE (e.g., 1996:39, 116-118). Furthermore, such an extended allolexy on the level of presumed semantic primitives is somewhat disturbing. Finally, by combining PERSON with the indefinite SOMEONE and interrogative WHO, on the one hand, and THING with the indefinite SOMETHING and interrogative WHAT, on the other, Wierzbicka appears to mix up categorial and functional domains. Thus, SOMEONE, as Wierzbicka uses it in her definitions of more complex notions, can perfectly be defined in terms of the category PERSON (in order to select human referents from all the entities) and a couple of other of Wierzbicka’s own semantic primitives (in order to convey the idea of indefinite reference), e.g. as the person I do not know or the person I know but you do not, etc., depending on the kind of indefinite meaning intended. The same is true for WHO,¹³ which may also be decomposed into PERSON and an

¹¹ Wierzbicka (1996) herself does not provide an explicit definition of the concept PERSON via the concept PEOPLE.

¹² The same primes are also reproduced in Goddard (2001).

¹³ Wierzbicka (1996:38) argues that the “interrogative meaning is not an inherent part of the

interrogative component.

What is more, it appears that neither PERSON, nor PEOPLE, nor THING need to be semantic primitives. They can all be explained by means of other, more basic notions, such as I, THE SAME, ONE, ALL, LIKE, KIND (OF), NOT, which actually already make part of Wierzbicka's set of the putative semantic primitives. The most important concept here is the notion I (ego, self), because I is the basis of all human experience and conceptualization of the world. It is the distinction between I and NOT I that appears to provide *the* most fundamental form of human categorization.¹⁴ The notions PERSON and PEOPLE are just abstractions of a certain degree from the most concrete of all possible notions, viz. the notion I. This idea is surely not new. For instance, Janda (1996:325) argues within the framework of cognitive linguistics that "animacy is just one of the many distinctions that can be made along the scale of SELF vs. OTHER". An example of such a "SELF-OTHER continuum" is illustrated in Figure 1. It should be kept in mind, though, that as Janda (1996:326) correctly points out, "[t]he SELF is not a literal self, but an idealization of this concept". Indeed, the SELF, or rather I, is not a psychological but a linguistic concept here, and human conceptualization as embedded in the language is clearly a "consensus" product formed by the interaction between many individual Is. It is this idealized I that provides the basis for Wierzbicka's "stable irreducible core" of the notion PERSON "across all languages and cultures".

I would venture to suppose that such a prototypical person must be an average healthy adult human (not elderly and perhaps male),¹⁵ belonging to the same PEOPLE, i.e. the same autoidentified social group as the speakers of language X themselves. Wierzbicka (1996:40) dismisses the "well-known fact that in many languages, the word for 'people' is also used as a tribal name (as is

words *who* and *what* as such", because for instance, they can be "used in so-called 'embedded questions' [as] 'I know (don't know) who did it'". However, this cannot be accepted as a valid argument, because here Wierzbicka mixes up two different constructions and, apparently, allows herself be influenced by the formal identity between the interrogative and relative pronominals in English (and quite a few other languages).

¹⁴ "The fact of our physical embodiment gives a very definite character to our perceptual experience. Our world radiates out from *our bodies* as perceptual centers from which we see, hear, touch, taste, and smell our world. Our perceptual space defines a domain of macroscopic objects that reside at varying distances from us. From our central vantage point we can focus our attention on one object or perceptual field after another as we scan our world [...] At a certain distance from this perceptual center our world "fades off" into a perceptual horizon which no longer presents us with discrete objects" (Johnson 1987:124 via Janda 1996:325).

¹⁵ Other factors, that may have some relevance here, are, for instance, whether the individual is married and/or has engendered a child and is still capable of engendering children.

Figure 1. “The barest default contours of the SELF-OTHER continuum” (Janda 1996:326)¹⁶

SELF/FIGURE	>	>		
SELF	> HUMANS LIKE SELF	> HUMANS NOT LIKE SELF	> ANIMALS	> SMALL, DISCRETE, COUNTABLE CONCRETE OBJECTS
		>	>	> OTHER/GROUND
> OTHER	> PARTS OF OBJECTS	> MASSES AND COLLECTIVES	> LANDSCAPE FEATURES	> AMBIENT INTANGIBLES AND ABSTRACTIONS

pointed out by Greenberg 1966a:26”, regarding it as a “clear [...] case of polysemy, comparable to the polysemy of the English word *man* (1) a male human being, (2) a human being”. However, at the same time, she speaks of PEOPLE as “a social, rather than biological, category” (Wierzbicka 1996:41). In other words, by dismissing the “well-known fact” just mentioned she at the same time forgets that it is exactly because PEOPLE is a social and not biological category, it should not cover all “biologically” human beings. Rather, it is more natural for the category PEOPLE to apply to the members of one’s own society, which are real PEOPLE, other groups of humans being to various degrees just like PEOPLE. Using Wierzbicka’s semantic primitives, the notion PEOPLE can thus be defined as something like ‘all of the same kind as I (and like I)’ or ‘all the same as I’ and the notion of PERSON as something like ‘one of the same kind as I (and like I)’ or ‘one the same as I’ respectively.¹⁷ Such definitions of PERSON and PEOPLE are also very flexible and can be extended to many other non-prototypical PERSONS and PEOPLES as the understanding of what exactly counts as THE SAME or LIKE may vary. Finally, just as there are entities that are like PEOPLE, there are also entities that are not like PEOPLE, with various degrees of

¹⁶ “This exposition gives only the barest default contours of the SELF-OTHER continuum. It can be differently realized in different languages and is subject to speaker construal; although the trend is clear, it is not immutable” (Janda 1996:326).

¹⁷ Such a restricted understanding of PEOPLE would also be useful for many of Wierzbicka’s definitions of other less basic concepts. Thus, Wierzbicka’s definition of the English concept *bird* begins with the following frame: “people think things like this about creatures of this kind” (1996:163). However, it may be pointed out that not all people “think things like this about creatures of this kind”, only speakers of English do.

likeliness and unlikeliness possible. THING can thus be defined as the category including entities that are not like PEOPLE.

However, from a cross-linguistic perspective, not all THINGS are THINGS to the same extent. That is, just as in the case of PERSON, there is in all probability some “stable irreducible core” of the concept THING “across all languages and cultures”. Defining this core would require an in-depth study of its own, but I would guess that the following parameters are relevant. A prototypical THING is not alive, was not alive and cannot be alive, it does not have parts that are/were/can be alive. Neither can it be like living things in any respect, e.g. it should not be able to move or change its appearance in any way (at least not in a way that can be perceived immediately) without people directly acting upon it. It is the kind of thing that is seen and that can be felt and that by preference, can be acted upon in some way by people (e.g., can be hit, given, moved, etc.). Consequently, it is not the kind of thing that happens or is said, thought or done. Preferably, it is the kind of thing that can be made by people.

Given the fundamental nature of the distinction between PERSON and THING as a means of human categorization of entities, this distinction appears to be particularly suitable to be taken as the basis for a cross-linguistic comparison of interrogative pronominals. After all, pronominals are substitutes for nominals and nominals are basically linguistic devices for classifying and referring to entities. Admittedly, in particular languages more fine-grained distinctions may be made than just ‘who?’ vs. ‘what?’. However, if there is any distinction based on the categorization of entities, it will be minimally between ‘who?’ and ‘what?’. This does not mean, though, that a language needs to possess minimally two distinct lexemes, meaning ‘who?’ and ‘what?’ respectively (cf. Section III.9.2 for a discussion).

Although the distinction between PERSON and THING is deeply entrenched in the human categorization of entities, PERSON and THING are not semantic primitives. As has been discussed above, even though these concepts have some stable irreducible core across all languages and cultures, as such they are flexible and language-specific.¹⁸ That is, it is by no means obvious how the two concepts as such should be used for the purpose of cross-linguistic comparison in a consistent and straightforward way. Therefore, I will confine the notion PERSON to biological humans (either alive or dead) and the notion THING to concrete biologically non-living entities,¹⁹ which can be referred to with the usual label

¹⁸ Admittedly, the extent of language-specific variation with PERSON and THING is probably much smaller than with most other meanings.

¹⁹ “Non-living” here is not the same as dead, since dead presupposes that the entity used to be alive. Admittedly, there may be no universally accepted scientific definition of life. Still, there is much more consensus about what life is between scientists than there appears to be between the

(concrete) *inanimates*. Delimited this way, *THING* will more or less correspond to the core of the notion *THING* across languages suggested above. Biologically living entities, which can be summarily referred to with the usual label (*non-human*) *animates*, fall somewhere in between these two notions thus delimited. However, by default, they are *THINGS*. I also consider by default as *THINGS* things happening, things done, said and thought, as well as things that are felt but cannot be seen are also considered. This last group can be subsumed under the label *abstract (or non-concrete) entities*.

The situation is more complicated with various mythological concepts. Languages regularly conceptualize mythological entities as though they were real world entities and consequently, they apply to them the same distinctions as to real world entities. Moreover, some real world entities may at the same time be conceptualized as mythological entities, e.g. a star may be believed to be a god. By default, mythological entities are clearly *THINGS*, since they do not refer to biological humans.

2.5 Identity vs. classification and proper names vs. common nouns

A non-selective interrogative pronominal questioning the identity of a person, such as English *who?*, is a form basically aimed at obtaining a (personal) proper name as answer,²⁰ or at least some specific personal description, cf. (22).²¹

- (22) [Persons A and B see person X. Person B appears to be familiar with X. Person A asks:] *Who is this?*
- a. [B:] *It's John.*
 - b. [B:] *It's my brother/ my doctor.*
 - c. [B:] [?] *It's the doctor.*

systems of conceptualization of the world as embedded in different languages.

²⁰ Following Van Langendonck (1999:95, 2007:6), proper name is defined as lexeme “assigned to an *ad hoc* referent in an *ad hoc* name-giving act”. Van Langendonck (1999, 2007) further distinguishes between “proper names” and “proprial lemmas (lexemes)”. Proprial lemmas are lexicon entries that are typically used as proper names. Furthermore, some nominal lemmas, “which are called ‘appellative proper names’ by a few German onomasticians: names of brands, languages, diseases, and so on”, take an intermediate position between proprial and common status. These can be called “proprio-appellative lemmas” (Van Langendonck 2007:254).

²¹ A definite description as an answer would normally be possible when accompanied by some additional explanation, such as *...the one we talked about yesterday*, or if it is a so-called “monoreferential appellative expression” (cf. Van Langendonck 2007:102-106), such as *the moon*, *the son of my neighbours* (provided they have only one son) or *the (village) doctor* (in the context of a village).

- d. [B:] ?? *It's a doctor.*

Answer (22d) with an indefinite description would suggest that the question was not about the identity of the person, but probably about the person's classification (class membership). However, such an answer does not appear adequate in the context described for (22).

The situation is somewhat less trivial when the question is about a thing, viz. when English would use *what?*. In this respect, consider example (23).²²

- (23) [Persons A and B see thing X. Person B appears to be familiar with X. Person A asks:] *What is this?*
- a. [B:] *It's my boomerang/ my neighbour's cherry-tree.*
 - b. [B:] *It's a boomerang/ a cherry-tree.*
 - c. [B:] ?? *It's the boomerang/ the cherry-tree.*

What is remarkable here is that an indefinite description, as in (23b), would be just as normal an answer to the question *What is this?* as a specific description, as in (23a). In fact, an indefinite description may be even *more* typical here. In other words, non-selective thing-interrogatives, such as English *what?*, are primarily intended to ask for a classification of a thing, rather than its identity in the strict sense, because prototypically, things are referred to with descriptions (common nouns) and not with proper names.

Admittedly, some things, just like persons, may have proper names, as for instance pets and places normally do. However, a thing does not need to have a proper name and most things do not have it, while for a person, a proper name is an indispensable attribute and expression *par excellence* of the person's identity. The asymmetry between 'who?', primarily related to identification and proper names, and 'what?', primarily related to classification and common nouns, appears to reflect one of the basic principles of human conceptualization of the world, viz. things are normally conceptualized as members of a class of non-unique individuals, while persons are typically conceptualized as uniquely identifiable individuals.

2.6 Summary

I believe that for purposes of cross-linguistic comparison, 'who?' and 'what?' are best defined as idealizations resulting out of the interaction between several parameters within a single conceptual space, as represented in Figure 2.

²² The same remark about definite descriptions as answers as applies to (22) above is applicable here as well.

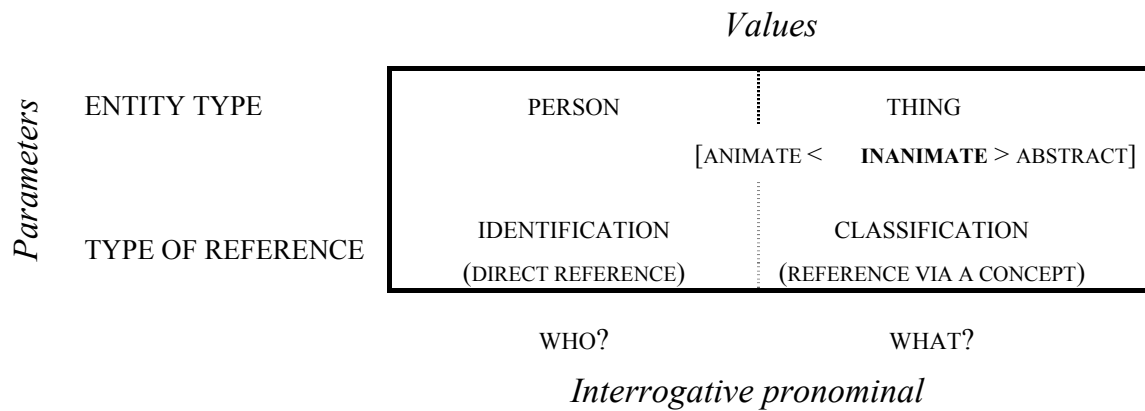
Figure 2. Conceptual space for delimiting the prototypical functions of non-selective interrogative pronominals

		<i>Values</i>	
<i>Parameters</i>	ENTITY TYPE	PERSON	THING
		[ANIMATE < INANIMATE > ABSTRACT]	
	TYPE OF REFERENCE	IDENTIFICATION (DIRECT REFERENCE)	CLASSIFICATION (REFERENCE VIA A CONCEPT)
	EXPECTED ANSWER	PROPER NAME	COMMON NOUN (DESCRIPTION, APPELLATIVE)
		WHO?	WHAT?
		<i>Interrogative pronominal</i>	

Figure 2 can be interpreted as follows. The prototypical interrogative WHO? is an interrogative pronominal asking for identification of a person (a biological human, cf. Section I.2.4) and expecting a proper name as an answer. The prototypical interrogative WHAT? is an interrogative pronominal asking for classification of a thing (any entity other than a biological human, but ideally a concrete biologically non-living entity, cf. Section I.2.4) and expecting a common noun as an answer.

Note that strictly speaking, the third parameter, viz. the expected answer, is to a certain extent redundant, because its values can be defined in terms of prototypical correlates of the respective values of the second parameter, the type of reference. Thus, a proper name, as an inherently referring nominal, is clearly the prototypical means of identifying a referent, whereas a common noun referring via a concept is the prototypical means of its classification. Therefore, in principle, the conceptual space presented in Figure 2 can be further reduced in the way illustrated in Figure 3. However, sometimes the expected answer does play an irreducible role in the choice of an interrogative pronominal. For instance, in English it appears to be possible to use *who* in questions about animals, more specifically domestic animals, only to ask about their proper names, such as *Fido*, but not their identification by means of some specific description (even a monoreferential one), such as *the dog of my neighbours*.

Figure 3. Conceptual space for delimiting the prototypical functions of non-selective interrogative pronominals (the reduced version)



In the remainder of the text, I may sometimes use the following shortcuts to refer to the definitions just discussed in a more compact way, cf. (24).

- (24) a. ‘who?’ → [person + identification (+ proper name)]
 b. ‘what?’ → [inanimate thing + classification (+ common noun)]

3 Research questions

The main question of the present study can be formulated as follows:

- (25) Is lexical differentiation between ‘who?’ and ‘what?’ universal?

Question (25) presupposes, however, that question (26) is answered first.

- (26) What qualifies as a lack of differentiation between ‘who?’ and ‘what?’?

If the answer to question (25) is positive, one may only ask oneself, why so? If it is negative, a whole lot of more specific questions can be considered, some of which are summarized in (27).

- (27) a. What kinds of lack of differentiation between ‘who?’ and ‘what?’ are attested?
 b. How widespread is the lack of differentiation between ‘who?’ and ‘what?’ in the languages of the world?
 c. Can any patterns be discerned in terms of genetic or areal distribution of the languages allowing for a lack of differentiation between ‘who?’ and ‘what?’?
 d. What could the origins (if any) of the lack of differentiation between ‘who?’ and ‘what?’ be in each particular case?
 e. What disambiguation strategies (if any) are used by the languages

lacking differentiation between ‘who?’ and ‘what?’?

- f. Why do (most) languages oblige their speakers to differentiate between ‘who?’ and ‘what?’? Or from a different angle: How is it possible that (some) languages do not oblige their speakers to differentiate between ‘who?’ and ‘what?’, given the fundamental nature of the distinction between persons and things as the basic means of human categorization of entities?

Obviously, many of these questions deserve an in-depth study of their own. Due to limitations of space and especially time, it will not be possible to answer all of them in the same degree of detail. Moreover, the information necessary to answer certain questions is lacking in most grammatical descriptions.

Admittedly, in part, the research questions outlined above have been formed by the definition of the object of study adopted (cf. Section I.2). Thus, one could have also taken as the main question the possible cut-off points within the systems of non-selective interrogative pronominals of the languages of the world. Moreover, my choice has been determined by the need to test the claim often found in typological literature that a lack of differentiation between ‘who?’ and ‘what?’ is either impossible (Wierzbicka 1996) or at least extremely marginal, with just a couple of exceptions attested (e.g., Hjelmslev 1956; Maytinskaya 1969; Ultan 1978; Lindström 1995; Siemund 2001; Bhat 2004). Given that a couple of cases of lack of differentiation between ‘who?’ and ‘what?’ have been mentioned in the typological literature, it is obviously difficult to maintain Wierzbicka’s (1996:38) claim that “[a]ll languages have words for WHO and WHAT, and can distinguish lexically between the questions ‘What is this?’ and ‘Who is this?’”. Consequently, question (25) is already largely answered. I say largely because it is desirable to control the reported cases of lack of differentiation with primary sources. Indeed, in some cases, this recheck proved to be justified. For instance, Ultan (1978:229) mentions Sango (Creole; Central African Republic) and Maytinskaya (1969:221, via Maria Koptjevskaja-Tamm, p.c.) Ewe (Niger-Congo, Kwa; Ghana) as languages lacking differentiation between ‘who?’ and ‘what?’, but this is not substantiated by the sources (Samarin 1967:74-75, 217; Thornell 1997:76; Madeleine Somte, p.c., for Sango; Pasch 1995:79; Westermann 1961:32 for Ewe).

4 What qualifies as a lack of differentiation between ‘who?’ and ‘what?’ (and what does not)

4.1 Simple cases

What qualifies as a genuine lack of differentiation between ‘who?’ and ‘what?’? Obviously, the language does not differentiate between ‘who?’ and ‘what?’, if it

has only one non-selective interrogative pronominal meaning either ‘who?’ or ‘what?’ depending on the linguistic or extralinguistic context. For instance, Urarina, a linguistic isolate spoken in northern Peru, has only one general interrogative pronominal *dʒa* ‘who?, what?’, as illustrated in (28).

Urarina

- (28) a. *dʒa kwara-i tahia?*
 IPW see-2 over.there
 ‘Whom/what have you seen over there?’ (Olawsky 2006:816)
- b. *dʒa turu-e?*
 IPW arrive-3_E
 ‘Who has come?’ or possibly ‘What has arrived?’ (Olawsky 2006:815)
- c. *dʒa kurete-i na-ĩ baha-anu?*
 IPW buy-2 say-PTCP ask-1SG_A
 ‘“What have you bought?”, I asked him’ (Olawsky 2006:816)

4.2 Complications

However, most of the world’s languages are not like Urarina, which makes the differentiation between ‘who?’ and ‘what?’ or lack thereof a much less obvious issue. To begin with, most languages have more than one non-selective interrogative pronominal and even though in most cases there are just two such pronominals, the distribution in the patterns of use between the two may differ substantially from language to language. Non-selective interrogative pronominals may vary in gender, number and/or case. Furthermore, if there is a lack of differentiation between ‘who?’ and ‘what?’ in a given language, it does not have to be obligatory. Languages may also just *allow* the speakers to leave the distinction between ‘who?’ and ‘what?’ undifferentiated without obliging them to do so. Besides, there are situations when speakers do not have the slightest idea of whether the entity in question is a person or a thing (or for one or another reason, do not wish to make any conjecture on this point), but they still need to choose one of the non-selective interrogative pronominals available. Finally, languages may use non-selective interrogative pronominals in syntactic constructions, which at first sight, may look very much like questions about identification or classification, but in reality are not.

Based on the definitions of ‘who?’ and ‘what?’ formulated in Section 1.2.6, in general terms the lack of differentiation between ‘who?’ and ‘what?’ can be defined as the possibility to use one and the same interrogative pronominal for both ‘who?’ [person + identification (+ proper name)] and ‘what?’ [thing + classification (+ common noun)]. However, for reasons such as those mentioned above, in particular languages the situation may become rather complicated. In

what follows, I will discuss some typical complications encountered in the languages of the world.

4.2.1 Syntactic function: predicate vs. non-predicate and case

There is a fundamental difference between the predicative and non-predicative uses of nominals (and pronominals), which corresponds to the fundamental distinction between predicates and terms in the propositional structure of a clause. Given this, the two types of the use of interrogative pronominals should be considered apart regarding a lack or presence of differentiation between ‘who?’ and ‘what?’. That is, if a language allows for a lack of differentiation between ‘who?’ and ‘what?’ only with predicatively used non-selective interrogative pronominals (or only in non-predicative use respectively),²³ such a lack of differentiation deserves to be described on equal terms with the full lack of differentiation between ‘who?’ and ‘what?’ in languages like Urarina, as presented above (Section I.4.1).

The lack of differentiation between ‘who?’ and ‘what?’ in the predicative function only can be illustrated with the example of the Modern French interrogative *quel* (29), which is otherwise used attributively as ‘which [N]?, what (kind of) [N]?’.

French

- (29) a. *quel est cet arbre?*
 IPW is DEM tree
 ‘What tree is this/that? (lit.: ‘What is this/that tree?’)’
 b. *quel est cet homme?*
 IPW is DEM man
 ‘Who is this/that man?’

Note that French only allows for a lack of differentiation between ‘who?’ and ‘what?’ in the predicate, but does not oblige the speakers to do so. See Section III.3.1.3.1 for more details. Consider also Old English discussed in Section III.3.1.4 and Ahaggar Tuareg (Afro-Asiatic, Berber; Algeria) in Section III.2.3.2.2.2. It may be worth remarking here that when a given language differentiates between ‘who?’ and ‘what?’ in non-predicative use but allows for a lack of differentiation between ‘who?’ and ‘what?’ in predicative use, the latter possibility is hardly ever mentioned explicitly in the sources. As a result, most of such cases have not found their way into the present study.

²³ Running a bit ahead, it may be worth mentioning that there appears to be an implicational relation here. That is, if ‘who?’ and ‘what?’ must be distinguished in non-predicatively used non-selective interrogative pronominals, then also in predicative use.

Non-predicative syntactic functions are usually marked in combination with various semantic roles. Morphologically, such markers may be bound, in which case one usually speaks about case, or non-bound, in which case one usually speaks about adpositions.^{24, 25} Some cases (or adpositions), such as genitive, instrumental, or locative, usually have a much more prominent semantic component to them than the others, such as nominative, accusative, ergative, or absolutive. Accordingly, the first cases are sometimes called semantic and the latter syntactic. Neutralization of the distinction between ‘who?’ and ‘what?’ in certain parts of the case paradigm is not uncommon in the languages of the world, as can be illustrated with the example of Urdu in Table 1, where the distinction between ‘who?’ and ‘what?’ is neutralized in all cases except the nominative/absolutive.²⁶

Table 1. Interrogative pronominals ‘who?’ and ‘what?’ in Urdu (Indo-European, Indo-Iranian; Pakistan; Dymshic 1962:54)

	SG		PL	
	‘who?’	‘what?’	‘who?’	‘what?’
NOM/ABS	<i>kaun</i>	<i>kyā</i>	<i>kaun</i>	<i>kyā</i>
ERG	<i>kis ne</i>		<i>kīnhoñ ne</i>	
OBL	<i>kis</i>		<i>kīn</i>	
DAT/ACC	<i>kise</i>		<i>kīnheñ</i>	

I do not consider such a partial neutralization in case as a lack of differentiation between ‘who?’ and ‘what?’. However, particular attention is necessary when the distinction between ‘who?’ and ‘what?’ appears to be neutralized in syntactic cases, but not in (some) semantic cases. The reason is that semantic cases tend to impose various restrictions on the possible semantics of the nominal involved. For instance, instrumental may be used only with inanimate nominals and dative only with human ones. What is more, semantic cases appear to be able to select

²⁴ Usually, there are also other distinctions between adpositions and case markers, but these are of little relevance here.

²⁵ In a language with morphological case, predicatively used nominals are marked for case as well. However, this does not take away the fundamental difference between the predicative and non-predicative uses of nominals (and pronominals). That in a given language a nominal must be morphologically marked for case in both non-predicative and predicative functions demonstrates only that morphological case is a grammatical category of nominals in this language.

²⁶ Urdu is characterized by the so-called “split-ergativity”, with the split being conditioned by tense-aspect (cf. Dymshic 1962:111).

one of the meanings of a given form in accordance with their most typical use. For instance, in German the genitive of both *wer* ‘who?’ and *was* ‘what?’ is *wessen*. However, *wessen* is only very rarely used as the genitive of ‘what?’ and never attributively (cf. Nau 1999:136-138, 141), which appears to be due to the prominent role of the possessor marking function in the semantics of the genitive.

We now can imagine a language that for instance next to a dedicated ‘what?’ interrogative, also has another interrogative, which in the nominative and accusative can mean both ‘who?’ and ‘what?’, but which only means ‘who?’ in the genitive. All other things being equal, I consider such an interrogative as a genuine instance of lack of differentiation between ‘who?’ and ‘what?’, because the human meaning ‘who?’ of such an interrogative in the genitive may be accounted for by the prominent role of the possessor marking function in the semantics of the genitive.

4.2.2 Gender and number

Besides case, interrogative pronominals may also vary in gender and/or number. If a given gender or number form of the non-selective interrogative pronominal can be used as both ‘who?’ and ‘what?’, I consider such a form as an instance of a lack of differentiation between ‘who?’ and ‘what?’ on a par with the full lack of differentiation between ‘who?’ and ‘what?’ as presented in Section I.4.1.²⁷ The reason is that unlike case, which is a structural feature assigned to the lexical form within a given clause, gender and number are, as it were, lexeme-internal features. This is particularly clear with number, because a different number form normally reflects a different number of entities of the same kind in the world. Thus, the dual form denotes two entities of the same kind, while singular denotes just one such entity. With gender, the relation to reality is less direct, viz. through the human conceptualization of entities and the classification of concepts in language. Thus, as a rule, nominals based on the same root or even having the same form but classified as belonging to two different genders are also conceptualized as referring to different kinds of entities.

The lack of differentiation between ‘who?’ and ‘what?’ involving gender can be demonstrated with the example of Apurinã (Arawakan, Brazil). In Apurinã, the masculine interrogative pronominal *ke-ru-pa* appears to be the default form in questions about humans and consequently the closest equivalent of ‘who?’ (30a). However, it can also be used as ‘what?’ (30b), since the masculine is also the default gender for animate and inanimate things. At the same time, the feminine

²⁷ Note, however, that a reservation may need to be made here with respect to non-human animates, i.e. biologically living entities (and those mythological entities conceptualized as live in a given language), as will be discussed in Section I.4.2.3.2.

form *ke-ro-pa* appears to be restricted to questions about women (possibly, also female animals, but this is not specified in the source; cf. Section III.8.2.3).

Apurinã (non-North Arawakan; Brazil)

- (30) a. *ke-ru-pa* *umaka?*
 IPW-3M-Q sleep
 ‘Who sleeps?’ (da Silva Facundes 2000:365)
- b. *u-kara* *ie* *ke-ru-pa?* [*pūtanuru pitximunanhi*]
 3M-DISTAL then IPW-3M-Q
 ‘[A:] What is it?! [B: ‘The penis body of your husband!]]’ (da Silva Facundes 2000:491)
- c. *ke-ro-pa?*
 IPW-3F-Q
 ‘Who is she?’ (da Silva Facundes 2000:123)

Somewhat similarly, in Latin the masculine singular interrogative pronominal *qui(s)* is the default form in non-selective questions about humans and consequently the closest equivalent of ‘who?’. However, at least predicatively, it also seems possible to use it as ‘what?’ in non-selective questions about things referred to with masculine nominals, such as *piscis* ‘fish’ (31a) or *fluvius* ‘river’ (31b) (cf. Section III.3.1.6).

Latin²⁸

- (31) a. *qui(s)* *ille* *piscis* *est?*
 IPW.M.SG.NOM DEM.M.SG.NOM fish.M.SG.NOM is
 ‘What is that fish?’
- b. *qui(s)* *ille* *fluvius* *est?*
 IPW.M.SG.NOM DEM.M.SG.NOM river.M.SG.NOM is
 ‘What is that river?’

In Ancient Greek, the situation is quite similar, but in addition, Ancient Greek does not distinguish ‘who?’ and ‘what?’ in the dual, as illustrated in Table 2.

4.2.3 Non-prototypical combinations of values (with respect to the choice of a non-selective interrogative pronominal)

4.2.3.1 A non-prototypical combination of values vs. lack of differentiation

Based on the definitions of ‘who?’ and ‘what?’ formulated in Section I.2.6, I have defined the lack of differentiation between ‘who?’ and ‘what?’ as the

²⁸ According to Marc Van Uytendange, Dirk Sacré and Pierre Swiggers (p.c.), these constructed examples represent correct Classical Latin.

possibility to use one and the same interrogative pronominal for both ‘who?’ [person + identification (+ proper name)] and ‘what?’ [thing + classification (+

Table 2. Paradigm of the Ancient Greek interrogative *tis* (Kühner 1890-1904:I, §176)

	SG		PL		DU
	M (& F)	N	M (& F)	N	M (& F), N
NOM	<i>tis</i>	<i>ti</i>	<i>tines</i>	<i>tina</i>	<i>tine</i>
ACC	<i>tina</i>		<i>tinas</i>		
GEN	<i>tinós</i>		<i>tinôn</i>		<i>tinoin</i>
DAT	<i>tini</i>		<i>tisi(n)</i>		

common noun)]. This has to be distinguished from *dominance* of one non-selective interrogative pronominal with respect to the other in cases of *non-prototypical combinations of values* on the parameters ENTITY TYPE and TYPE OF REFERENCE (& EXPECTED ANSWER), i.e. [person + classification (+ common noun)] and [thing + identification (+ proper name)] (as well as [animate thing + classification (+ common noun)], which is a special case discussed in Section I.4.2.3.2). For instance, in Russian the non-selective interrogative pronominal *kto* ‘who?’ can be used not only for [person + identification], as in (32a), but also for [person + classification], as in (32b), thus extending into the semantic domain prototypically covered by *čto* ‘what?’, at least on the parameter TYPE OF REFERENCE.

Russian

- (32) a. *A on kto voobščē?* [Da eto Petja. Ty ego ne znaeš’].
 and he who actually
 ‘[A:] Who is he actually? [B: Oh, it’s Pete. You don’t know him.]’
 b. *A on kto voobščē? Doktor?*
 and he who actually doctor
 ‘What is he actually? A doctor?’

Conversely, the translation of (32b) shows dominance of English *what* ‘what?’ in the same situation. That is, languages may prefer different strategies in cases of non-prototypical combinations of values.²⁹ Thus, in this particular case, some

²⁹ It should be emphasized that the use of the modifier *non-prototypical* in describing the combinations of values [person + classification (+ common noun)] and [thing + identification (+ proper name)] does not entail at all that a question for classification of a person or a question for identification of a thing is unnatural in any sense. What it only means is that the respective

languages, including Russian, will use ‘who?’ giving preference to the kind of entity questioned about, some, such as English, will use ‘what?’ giving preference to the type of reference involved. Still other languages prefer an avoidance strategy. Of course, avoidance strategies are available to all languages, including Russian and English, but some languages may make more extensive use of them than others. Usually, the preference is determined by the larger context. Some possible ways to avoid the use of ‘who?’ and ‘what?’ in the case of a non-prototypical combination of values are illustrated with example (33) from Taiwanese Min Nan Chinese, where neither *ziǎ* ‘who?’ nor *sā̃/s(i)ã-mì* ‘what?’ are possible, and example (34) from Ancient Greek, where the masculine plural nominative form *poioi* of the otherwise usually attributive interrogative *poios* ‘what (kind of) [N]?, which [N]?’ is used instead of the masculine³⁰ plural nominative form *tines* of the regular interrogative pronominal *tis*.

Taiwanese Min Nan Chinese (Sino-Tibetan; Taiwan; Cheng-Fu Chen, p.c.)

(33) *Zi-hūi-guaⁿ le?*

commander PTCL

‘[A: I joined the army yesterday. Half of the officers are men, the others are women. B:] And what is the commander? (lit.: ‘(What about) the commander?’) [A: It’s a man.]

Ancient Greek

(34) *poioi* *k’* *eit’* *Oduṣēi*

IPW.M.PL.NOM PTCL be.PRS.OPT.ACT.2PL PROP.M.DAT.SG

amunemen

keep.off.PRS.OPT.INF.ACT

‘What are you to keep off Odusseus?’ (Liddell *et al.* 1940 under *poios*)

Note that a given language may have a lack of differentiation between ‘who?’ and ‘what?’ and at the same time use ‘what?’ in non-prototypical combinations of values involving persons and/or ‘who?’ in non-prototypical combinations of values involving things. Of course, this is possible only if the language has more than one interrogative pronominal. Thus, French, as mentioned in Section I.4.2.1, allows for a lack of differentiation between ‘who?’ and ‘what?’ with a predicatively used interrogative *quel* (cf. Section III.3.1.3.1 for more details). At the same time, French has a human *qui* ‘who?’ and a non-

combinations of values are not prototypically associated with one of the two non-selective interrogative pronominals, viz. ‘who?’ or ‘what?’. This is also why different languages distinguishing between ‘who?’ and ‘what?’ may opt for a different non-selective interrogative pronominal in cases of such non-prototypical combinations of values.

³⁰ In fact, non-neuter, because the masculine and the feminine are not distinguished here.

human *que* or *quoi* ‘what?’ and ‘what?’ is used to in the case of a non-prototypical combination of values of the kind [person + classification (+ common noun)], as in (35).

French

- (35) [A:] *t’es avocat?* [B:] *non* [A:] *t’es docteur?* [B:] *non*
 you-are lawyer no you-are doctor no
 [A:] *t’es quoi alors?* [B:] *je suis linguiste*
 you-are what then I am linguist
 ‘[A:] Are you a lawyer? [B:] No. [A:] Are you a doctor? [B:] No. [A:]
 What are you then? [B:] I’m a linguist’

4.2.3.2 Non-human animates: a border case

Some kinds of uses of non-selective interrogative pronominals may be very close to the lack of differentiation between ‘who?’ and ‘what?’, such as the use of Russian *kto* ‘who?’ in questions about animals (36).

Russian

- (36) a. *A eto kto? Korova?*
 and this who cow
 ‘And what is this? A cow?’
 b. *Kto mog ostavit’ takie sledy? Los’?*
 who could leave such traces elk
 ‘[In the woods, looking at fresh hoofprints on the ground:] What could
 have left such tracks? An elk?’

I prefer to consider this and similar uses of otherwise human non-selective interrogative pronominals prototypically used for [person + identification (+ proper name)] in questions about animate things [animate thing + classification (+ common noun)] as an instance of a non-prototypical combination of values rather than a lack of differentiation. As discussed in Section I.2.4, even though by default (non-human) animates, viz. biologically living entities, are things, strictly speaking they fall somewhere in between persons and prototypical things, viz. inanimate things (as delimited in Section I.2.4). Thus, many languages treat non-human animates similarly to humans in various respects, although these languages may differ as to what kinds of animates they assimilate to humans. At the same time, in questions about non-human animates, many languages tend to avoid using both ‘who?’ and ‘what?’, especially when it is explicitly mentioned within the same predication that the entity in question is a non-human animate. For instance, although Russian allows the use of *kto* ‘who?’ in (36), *kto* ‘who?’ sounds very awkward in contexts such as (37) and *čto* ‘what?’ is not possible at

all. Rather, the construction as in (38) should be used.

Russian

- (37) a. *ʔKto eto životnoe?*
 who this animal
 b. **Čto eto životnoe?*
 what this animal
 ‘What is this animal?’
- (38) *Čto eto za životnoe?*
 what this for animal
 ‘What animal is this? (lit.: ‘What (is) this for animal?’)

In French, as noted by Riegel *et al.* (2001:394), it is somewhat difficult to ask a question about an animal using the basic non-selective interrogative pronominals *qui* ‘who?’ or *que* (or *quoi*) ‘what?’, since *qui* “oriented to humans” (“orienté vers l’humain”) and *que* “oriented to inanimates” (“orienté vers le non-animé”) “do not fit very well” (“ne conviennent pas exactement”). Similarly, in Danish, according to Allan *et al.* (2003:195), “since *hvem* [‘who?’] is used for human beings and *hvad* [‘what?’] for inanimate objects, there is a gap in the semantic field covered by these two interrogative pronouns which excludes animals. It is thus equally ungrammatical to use *hvem* and *hvad* in” the context of (39), instead *hvad for et* ‘what kind of one’ (lit. ‘what for one?’) (39c) needs to be used.

Danish (Allan *et al.* 2003:195)

- (39) [*Et dyr må have lavet disse spor.*]
 a. **Hvem er det?*
 who is it
 b. **Hvad er det?*
 what is it
 c. *Hvad er det for et?*
 what is it for one
 ‘[An animal must have made these tracks.] What is it?’

Admittedly, my decision to count examples like (36) above as a non-prototypical combination of values rather than a lack of differentiation between ‘who?’ and ‘what?’ may be somewhat arbitrary. In fact, given that as a rule animate things can be rather straightforwardly distinguished from persons on a language-external basis, it may have been much easier in many cases to simply classify such examples as a lack of differentiation between ‘who?’ and ‘what?’. However, I believe that this would have been too much of an oversimplification.³¹

³¹ Next to animates, languages may also assimilate various inanimates, viz. biologically non-live

4.2.3.3 Neither a lack of differentiation nor a non-prototypical combination of values

4.2.3.3.1 “No conjecture”-contexts

There are situations when speakers do not have the slightest idea of whether the entity in question is a person or a thing (or for one or another reason, do not wish to make any conjecture on this point), but they still need to choose one of the non-selective interrogative pronominals that the language makes available to them, as in (40).

- (40) [To a person standing at the window and watching outside all the time:]
What are you looking at there?

Such contexts, which are clearly different from both the lack of differentiation between ‘who?’ and ‘what?’ and any kind of non-prototypical combination of values, may be called something like “*no conjecture*”-contexts.

In a “no conjecture”-context, the languages of the world appear to be just like English in preferring the interrogative ‘what?’ rather than ‘who?’. In all probability, this preference is due to the fact that semantically the category **THING** is clearly less marked, less salient than the category **PERSON** (cf. Section I.2.4). I know of only two near exceptions to this universal, viz. Wambaya (Mirndi group) and Gurr-goni (Maningrida), two non-Pama-Nyungan languages of northern Australia. The non-selective interrogative pronominals in these two languages vary in gender (and number) rather than humanness vs. non-humanness. The masculine interrogative is the default form in questions about humans and thus the closest equivalent of ‘who?’. However, the masculine form is also used in “no conjecture”-contexts because of the default status of the masculine agreement pattern in these two languages (cf. Section III.6.3.3 for more details).³² I consider these two cases to be only *near* exceptions, because

entities, to humans. Usually, nominals referring to such inanimate things are described as grammatically animate. Some languages may use the same interrogative in questions about them as they would use in questions about humans or other animate things. As a rule, I prefer to treat such cases as a lack of differentiation between ‘who?’ and ‘what?’ when (grammatically animate) inanimate things are involved and as a non-prototypical combination of values when (grammatically animate) animate things are involved (cf. also Section II.4.2.2).

³² Note that cases like Apurinã (Arawakan; Brazil; Section III.8.2.3), mentioned in Section I.4.2.2, are rather different. In Apurinã, the masculine interrogative pronominal *ke-ru-pa* appears to be the default option for ‘who?’. It is not clear from the description of Apurinã by da Silva Facundes (2000), whether it can also be used in “no conjecture”-contexts. However, even if this is the case, Apurinã still differs from Wambaya and Gurr-goni in two important respects. First, Apurinã has an interrogative pronominal form unmarked for gender, viz. *ke-pa/ki-pa*,

strictly speaking, the masculine interrogative is not a ‘who?’ interrogative in the same sense as, for instance, English *who* is.

4.2.3.3.2 Groups of people

Collective nominals denoting groups of people, such as *nation*, *team* or *family*, deserve particular attention with respect to the use of interrogative pronominals, as compared to plurals of non-collective personal nominals, such as *boys* or *recruits*. While plurals of non-collective personal nominals normally behave similarly to their singulars, there appears to be a difference between personal collectives that denote groups of people usually referred to by a proper name, such as ‘team’ or ‘family’, and those that are not, such as ‘committee’ or ‘squad’. In Russian, for instance, nominals of the former kind appear to strongly prefer *kto* ‘who?’ (41), while the latter tend to prefer *čto* ‘what?’ (42).³³

Russian

- (41) a. *Ja boleju za “Spartak”, a ty za kogo* čto?*
 I be.sick for PROP and you for who/what
 ‘I am a fan of Spartak (a football team), and what do you support? (lit.: ‘...and you of whom?’)’
- b. [*V prošlom godu kubok Rossii vyjgral “Spartak”.*]
*A kto/*čto ego vyjgraet v etom godu?*
 and who/what it will.win in this year
 ‘[Last year, it was “Spartak” who won the Cup of Russia.] And who is going to win it this year?’

which would be even better suited for “no conjecture”-contexts. Second and most importantly, the masculine form *ke-ru-pa* is also used when the speaker is clearly aware that the referent is non-human, as in (30b). The only requirement seems to be that the referent is not a woman (or perhaps more broadly, not a female animate). That is, *ke-ru-pa* is both ‘who?’ and ‘what?’, so that its presumed use in “no conjecture”-contexts would not contradict the universal preference for ‘what?’ in such contexts. At the same time, in Wambaya and Gurr-goni, the masculine form, also functioning as the default option for the meaning ‘who?’, would not be used when the speaker is clearly aware that the referent at issue is non-human (or that the noun used to refer to is not masculine).

³³ Note that the choice between *kto* ‘who?’ and *čto* ‘what?’ here does not have to do anything with animacy, because personal collectives, such as *komanda* ‘team’, *sem’ja* ‘family’, *narod* ‘people (nation)’, *vzvod* ‘platoon’, as well as the football team name *Spartak*, are grammatically inanimate (even though the personal proper name *Spartak* is grammatically animate).

- (42) *Čem/*kem ty komandueš'? Vzvodom, rotoj, batareej?...*
 what.INS/who.INS you command platoon.INS company.INS battery.INS
 'What are you in command of? A platoon, a company, a battery?...' (based
 on Barulin 1980:38)

However, the division is not neat and many complications are possible. To begin with, according to Barulin (1980:38), it is often impossible to ask about personal collectives using either *kto* 'who?' or *čto* 'what?'.³⁴ Furthermore, to a large extent the choice between *kto* 'who?' or *čto* 'what?' appears to depend on the semantics of the predicate, viz. whether it typically selects human (or animate) arguments or non-human (or inanimate) ones. Finally, it is important to keep in mind that speakers either (i) may give more prominence to the fact that a given group is composed of humans, which would entail that the collective nominal referring to such a group denotes a kind of people, or (ii) they may perceive the group as an entity in its own right without relation to the fact that it is composed of humans, which would entail that the collective nominal referring to such a group denotes a kind of thing.³⁵ Therefore, it is probably most reasonable to consider the use of 'what?' in questions about personal collectives as neither a lack of differentiation between 'who?' and 'what?' nor a non-prototypical combination of values.

4.2.3.3.3 Deceptive appearances

Languages may use non-selective interrogative pronominals in syntactic constructions, which at first sight may look very much like questions about identification or classification, but in reality are not. Some cases, such as (43) from Dutch, may be more obvious than others, such as (44) from Russian.

Dutch

- (43) *Wat zijn we vandaag?*
 what are we today
 'What day is it today? (lit.: 'What are we today?')

Russian

- (44) *Kto vy? Čto vy?*
 who you what you
 '[Having finished talking with someone else, an official asks the person

³⁴ It should be pointed out, though, that most of the examples Barulin (1980:38) cites sound quite normal to me.

³⁵ To a certain extent, the same distinction appears to be reflected in the proper name-parameter mentioned above.

who approaches him and clearly wants something from him:] Who are you? What's your problem? (or 'What's up with you?'; lit.: 'What (are) you?')

The conventionalized meaning of the Dutch question in (43) is very different from what its literal translation may suggest, although the same question would also be used in translating English examples such as (45).

English

- (45) *So, what are we today, Dean? Are we rock stars, army rangers?* (from the series *Supernatural*, <http://community.tvguide.com/blog/Snfans-Blog/700010435>, retrieved 23.04.2007)

The situation with the Russian example in (44) is substantially different. Unlike in English or Dutch, it is not possible to use *čto* 'what?' in Russian in questions about classification of persons. The use of *čto* in the meaning 'what's up?' is possible only with pronominal subjects. It appears to derive from its use in the meaning 'why?' in sentences like (46).

Russian

- (46) *Nu, čto vy pereživaete? Ničego tut strašnogo...*
 come.on what you worry of.nothing here of.frightful
 '[Trying to help someone in distress to calm down:] Come on, why are you worrying? There's nothing to worry about here...'

Russian is not unique among the world's languages in allowing for the use of 'what?' in the construction 'what is person X?' with the meaning 'what's up/what's the matter with person X?' (or the like). In this respect, consider the discussion of Biblical Hebrew in Section III.2.3.4.1.3.3 and Ancient Egyptian in Section III.2.3.3.2.3, as well as the following examples from Thai (47), Kuuk Thaayorre (48) and Badaga (49).

Thai (Tai-Kadai; Thailand; David Smyth, p.c.; Pittayawat Pittayaporn, p.c.)³⁶

- (47) a. *John bpen a-rai?*
 John is what
 'What's up with John?', 'What's wrong with John?', 'What (disease) has John got?' (lit.: 'What is John?')
- b. *kăirn bpen a-rai?*
 arm is what
 '[A question to someone who walks into a room with a swelling on his arm, a wound, etc.:] What's (the matter) with your arm?' (lit.: 'What is

³⁶ The Thai examples here are transliterations, as provided by David Smyth.

(your) arm?’

Kuuk Thaayorre (Pama-Nyungan; Australia; Alice Gaby, p.c.)

(48) *John ngan?*

John what

‘What is John up to?’, ‘Why is John behaving as he is?’, ‘What’s the matter with John?’ (lit.: ‘What (is) John?’)

Badaga (Southern Dravidian; India; Christiane Pilot-Raichoor, p.c.)

(49) [A:] *Bīma ēna?* [B:] *Bīma buddu buṭṭa*

PROP what PROP fall.PTCP AUX.PFV.PST.3SG

‘[A:] What’s up with Bima? (What happened to Bima?) [B:] Bima has fallen’

The following examples from Kewa (50) and Korean (51) seem to be of the same kind, even though the exact meanings (and the associated connotations) appear to be somewhat different from the examples cited above, so that a direct translation with ‘what’s up with...?’ or ‘what’s the matter with...?’ would not go well here.

Kewa (Trans-New Guinea, Central and Western Main Section; Papua New Guinea; Karl Franklin, p.c.)

(50) *Yoaane-re ake ya-pae?*

John-TOP what is-Q

‘What kind of a person is this (particular) John?’ (‘suggesting that he is unusual in some way’)³⁷

Korean (Isolate; South Korea; Kyung-Ah Kim, p.c.)

(51) *ce salam mwe-ya?*

that person what-is.Q

‘What in the world is that person doing here?’ or ‘Who the heck is that person?’

These and similar constructions may be summarily referred to as something like ‘*what’s up?*’-constructions.

Another kind of construction that may be mentioned here can be illustrated with example (52) from Russian.

³⁷ This construction cannot be used to question the class membership of a given person in the strict sense, such as his or her ethnic origin, occupation, social position, sex, etc. Therefore, a somewhat more adequate translation might be something like ‘What’s the matter with this (particular) John?’.

Russian

- (52) *A čto takoe "staraja deva"?*
 and what such.N.SG.NOM old.F.SG.NOM maid.F.SG.NOM
 '[A: They say she is a spinster. B:] And what does *spinster* mean? (lit.: 'And what (is) as such *old maid*?')'

Here, the nominal *staraja deva* 'spinster, old maid' is used autonomously. It does not refer to anything but itself, viz. 'the word *spinster*', which in the Russian original is also made evident by the use of the neuter agreement pattern with *takoe* 'such', while *staraja deva* is feminine (as well as by the usual use of inverted commas with *staraja deva* in writing). Consequently, this is not a kind of non-prototypical combination of values. In fact, the English translation of (52) makes it particularly clear. Compare also a seemingly similar Russian example (53), which however represents an instance of a non-prototypical combination of values, as is also made clear by its English translation.

Russian

- (53) *A kto takaja staraja deva?*
 and who such.F.SG.NOM old.F.SG.NOM maid.F.SG.NOM
 "[A: They say she is a spinster. B:] And what is a spinster? (lit.: 'And who (is) as such an old maid?')"

Note that (53) differs from (52) not only in the use of *kto* 'who?' instead of *čto* 'what?', but also by the fact that *takaja* 'such' agrees in the feminine gender with *staraja deva*.

4.2.3.3.4 Special discourse forms

Some discourse forms, such as fairy tales and adult-to-infant talk, may differ from ordinary language use in the use of 'who?' and 'what?'. To begin with, there is a lot of anthropomorphization (or personification) in such discourse forms, which may result in the use of 'who?' in questions about animals in languages that otherwise do not allow such use. In English, for instance, questions about animals are normally formed with *what*, but in fairy tales and adult-to-infant talk such questions are often formed with *who* instead, as in (54).

English

- (54) [A question accompanying a picture of a heap of fallen leaves covering the ground:] *In the autumn... animals begin to hibernate in the leaves. Who do you think is under these leaves?* [The answer accompanied by a picture of a hedgehog:] *This hedgehog is sleeping under the leaves. It is too cold for him to hunt for food. Did you hear him snoring?* (from an online book "In the autumn...", <http://www.sebastianswan.org.uk>,

retrieved 13.04.2007)

Adults speaking to infants may also use a different strategy with respect to the choice of a non-selective interrogative pronominal in cases of non-prototypical combinations of values than they would apply in the ordinary language use, which has nothing to do with anthropomorphization. In Dutch, for instance, *wat* ‘what?’ is normally used in the case of a non-prototypical combination of values of the kind [person + classification (+ common noun)], as in (55a). However, in the case of adult-to-infant talk, *wie* ‘who?’ is often used instead, as in (55b). Note also that in the latter case, the answer will normally involve a definite article *de* rather than the indefinite one *een*.

Dutch

- (55) a. [A:] *Wat is dit?* [B:] *Eendokter*
 what is this a doctor
 ‘[A shows a picture of a doctor to B. B is not an infant. A:] What is this? [B:] A doctor.’
- b. [A:] *Wie is dit?* [B:] *De dokter*
 who is this the doctor
 ‘[A shows a picture of a doctor to B. B is an infant. A:] Who is this? [B:] The doctor.’

This difference between adult-to-infant talk and regular language use is very interesting as such. However, the present study is confined to regular language use.

5 Earlier work

Most grammatical descriptions of a reasonable length have a separate section on interrogative pronominals. However, as a rule, the interested reader will not find much more than forms of the interrogative pronominals provided with some broad glosses in the language of the publication and one or two examples of their use. Normally, a description would also mention when the interrogative pronominals vary in gender and case, although the paradigms are rarely provided in full. At the same time, if a given language has gender but the interrogatives pronominals do not carry any overt gender marking, the reader would rarely be informed explicitly about the agreement patterns that can be used with this or that interrogative pronominal. Similarly, whether the interrogative pronominals can inflect for number or not usually remains without mention.

When it comes to the semantics and patterns of use of the interrogative pronominals, most grammars (and dictionaries) happen to be satisfied with minimal descriptions. Often, such a description would consist of a statement of the kind “Language L has an animate interrogative X ‘who?’ and an inanimate

interrogative Y ‘what?’”. However, on closer examination, the terms *animate* and *inanimate* regularly prove to be used misleadingly for *human* and *non-human* respectively. It is also worth pointing out that the language the description is made in (especially, if it is the native language of the author) tends to strongly influence the way the semantics and patterns of use of the interrogative pronominals are described.³⁸ More recent grammatical descriptions are not necessarily better than the older ones.³⁹ In particular, descriptions made before the middle of the 20th century would often contain more information on the patterns of use of the interrogative pronominals, even though usually this information is more difficult to extract due to the regular lack of glosses (as well as transliterations in the case of languages traditionally using non-Latin-based scripts). The semantics and patterns of use of the interrogative pronominals of particular languages have rarely been the subject of special studies.⁴⁰ The papers in the collective volume edited by Chisholm *et al.* (1984) are also worth

³⁸ It is probably a commonplace, but people do tend to pay attention to, and consequently, mention explicitly only things that differ sufficiently from what is found in their own language (or the language they are writing in). For instance, a description made in English would rarely mention it when the language described allows to question about the classification of a person with a ‘what?’ interrogative (as in *What is John? A doctor?*). At the same time, such an English-based description would normally mention it when the language uses a construction literally meaning something like ‘Who is this man’s name?’ for ‘What is this man’s name?’. People also tend to recast systems found in other languages in terms of their own linguistic system. Thus, English-based descriptions often present interrogative pronominal systems based on gender in terms of *who* and *what* without paying attention to the semantics of genders and agreement patterns of the language at issue. This naturally results in rather inadequate descriptions. A somewhat different example can be found in Russian-based descriptions. Such a description may sometimes state something like “the interrogative pronoun X means both *kto* [‘who?’] and *čto* [‘what?’]”, when the language under description uses this interrogative pronominal both in questions about animals and inanimate things, because in Russian *kto* ‘who?’ applies to animals as well. Although such a statement is perfectly correct in Russian, it remains ambiguous unless further clarified in text or by means of examples. Furthermore, from a typological perspective, it can be rather misleading if taken over literally.

³⁹ Descriptions inspired by generative and various other “formal” frameworks are clearly the least useful for the kind of study undertaken here.

⁴⁰ However, cf. papers such as Bamgboṣe (1980) on Yoruba and Barulin (1980) on Russian. Furthermore, mostly for European languages, much information can often be found in (monolingual) academic grammars.

mentioning here, although they are much broader in their scope, with much less emphasis on semantic issues.

Typological literature considering interrogative pronominals from the point of view of their semantics and patterns of use rather than their syntax is virtually nonexistent. In fact, this is true for interrogative proforms in general. To the best of my knowledge, the only serious attempt to create a typology of interrogative proforms, with particular attention to the meanings distinguished and possible historical relations between the latter, has been made in two conference papers (2004a, 2004b) by Michael Cysouw. What one normally finds elsewhere is just a claim, made in passing, that all (Wierzbicka 1996) or nearly all languages (e.g., Hjelmslev 1956; Maytinskaya 1969; Ultan 1978; Lindström 1995; Siemund 2001; Bhat 2004), with perhaps a couple of exceptions, distinguish between ‘who?’ and ‘what?’. Of all the publications just mentioned, Lindström (1995) is the only paper dedicated exclusively to the examination of the latter claim. However, her sample, including just 24 languages, is obviously too small to allow her to arrive at any significant generalizations.⁴¹

Somewhat different questions are raised in papers by Mushin (1995) and Nau (1999). Thus, Mushin (1999) provides a very interesting account of interrogative and indefinite proforms, summarily referred to as “epistememes”, in 26 Australian languages. In particular, she examines “the grammatical properties, the categories of knowledge that are encoded, and the functions of the episteme class” (Mushin 1995:1). Nau (1999) discusses the interaction of ‘who?’ and ‘what?’ with morphological case in several European and a couple of Australian languages. Despite the very restricted nature of her sample, Nau (1999) arrives at some insightful generalizations.

6 The sample

The issue of sampling has received a lot of attention in the typological literature (cf., e.g., Bell 1978; Dryer 1989; Perkins 1989; Nichols 1992; Rijkhoff *et al.* 1993; Bybee *et al.* 1994; Rijkhoff & Bakker 1998; Maslova 2000; Cysouw 2005; Widmann & Bakker 2006; Bakker, to appear). Various sampling methods proposed in the literature have their advantages and disadvantages. However, what they all appear to have in common is that they strive for constructing a sample *representative* of the world’s languages. (Inevitably, this turns out to be possible only within the usual limitations imposed above all by availability of time and (good quality) data on the subject studied). Consequently, most discussion normally revolves around the issue of representativity of the sample

⁴¹ In fact, Lindström’s (1995) sample is much smaller than, for instance, that of Maytinskaya (1969) or Ultan (1978).

constructed with a given sampling procedure. However, representativity as such is a relative notion and is always a matter of degree. A sample may be only more or less representative with respect to the subject studied and the questions one has in mind.

The present study is primarily concerned with exploring the diversity of the natural spoken languages in the domain of the non-selective interrogative pronominals ‘who?’ and ‘what?’ as regards (i) the universality of their formal differentiation and (ii) the patterns of their functional differentiation in languages where they are distinguished formally. Note in this respect that (i) the possible exceptions to the universality of formal differentiation between ‘who?’ and ‘what?’ are generally assumed to be extremely rare, while (ii) the possible patterns of functional differentiation between ‘who?’ and ‘what?’ have never been the object of a typological study and it seems to be generally taken for granted that in all languages the division is strictly determined by the features person and thing. All this points at a large variety sample as the sample that suits best the purposes of the present study.

As was pointed out in Section I.5, this study has been seriously hampered by a lack of relevant data, especially as far as the functional differentiation is concerned. Therefore, my sample is inevitably very much a convenience sample, which in turn implies a certain areal and genetic bias. At times, the genetic bias in my sample may be rather strong since for some linguistic groups I have considered many closely related languages when the relevant data happened to be available. The two most extreme examples are represented by some 500 Bantu and some 450 Austronesian languages, mostly from the databases of Bastin *et al.* (1999)⁴² and Blust *et al.* (2006) respectively. Note that considering closely related languages may be useful in several respects. To begin with, in view of the fact that reliable diachronic data for most of the world’s languages are generally nonexistent, comparison of closely related languages may be a very good means of assessing the degree of diachronic stability of the phenomenon under investigation. Furthermore, it often proves indispensable to consider closely related languages when a diachronic explanation for a certain phenomenon is sought for (e.g., when possible origins of a lack of differentiation between ‘who?’ and ‘what?’ are investigated).

I have tried to counterbalance the genetic, areal and especially bibliographic biases by contacting specialists working directly on various languages and linguistic groups. Typically, I asked them to answer the questionnaire I composed specifically for the present study (cf. Section I.7). All in all, even if

⁴² I am grateful to Yvonne Bastin for granting me access to the unpublished database that was used for Bastin *et al.* (1999). I am also grateful to the Linguistic Service of the Royal Museum for Central Africa in Tervuren for making these data available to me in electronic format.

my sample may be far from being balanced, it is surely diverse enough to allow for meaningful cross-linguistic generalizations.

My global sample comprises around 1850 languages.⁴³ This number includes 1048 languages of my own database, which are listed in Appendix C together with the highest genetic grouping they are attributed to in the latest edition of the *Ethnologue* (Gordon 2005). The remaining languages come from Bastin *et al.*'s (1999) and Blust *et al.*'s (2006) databases already mentioned above. I have not yet fully incorporated these two databases in my own database. Therefore, no list of these languages is provided.⁴⁴

Finally, note that within my global sample several smaller subsamples can be distinguished as regards various kinds of non-prototypical combinations of values. These samples have been obtained through a non-controlled reduction of the global sample. In Section II.5 I specifically discuss the most important of these samples, which contains almost 200 languages for which I have unequivocal data on the three kinds of non-prototypical combinations of values investigated in this study.

7 Gathering the data: published sources and the questionnaire

The data used in this study comes from two kinds of sources. The information on the formal differentiation between 'who?' and 'what?' mostly comes from published sources.⁴⁵ The information on the semantics and patterns of use of the non-selective interrogative pronominals in the languages of the world (perforce) comes mostly from the data that I gathered by means of a questionnaire from specialists and native speakers of particular languages. I am deeply indebted to all the people who made time to share their knowledge with me.⁴⁶ Their names can be found in Appendix A together with the names of the respective languages.⁴⁷

My questionnaire exists in two versions, a "regular" and a "light" one, both translated in Russian (entirely), French (entirely), Spanish (partially) and Portuguese (partially). The two versions in English can be found in Appendix

⁴³ The term *language* here is used as a cover term for both languages and dialects.

⁴⁴ The only information the two databases provide are the forms of 'who?' and 'what?' in the respective languages. All the languages from these databases that have not yet been included in my database appear to score negatively on lack of differentiation between 'who?' and 'what?'.

⁴⁵ Under published sources I also subsume sources such as online databases.

⁴⁶ I would also like to thank all the people who reacted to my requests for information even though they could not help me further for one or another reason.

⁴⁷ Note that not all the questionnaires have been answered to the same degree of completeness.

B.1 and Appendix B.2 respectively. The light version has been introduced only somewhere midway in the data gathering process. Since then I usually offered both versions together. The main purpose of the introduction of the light version was to increase the response rate.⁴⁸ After all, some information is still better than no information.

Note that the regular version has been evolving in time, mostly in response to the answers provided by the specialists consulted. However, by and large, the differences between the different versions concern only some minor details. I abstained from modifying the questionnaire considerably in order to enhance the comparability of the data gathered for different languages. The version that can be found in Appendix B.1 represents the last “edition” of my questionnaire.

⁴⁸ I estimate the overall response rate to be roughly around 25%, of which probably some 15% represent the informative responses.

Evey: Who, who are you?

V: Who? *Who* is but the form following the function of *what...* and what I am is a man in a mask.

Evey: I can see that.

V: Of course you can. I'm not questioning your powers of observation, I'm merely remarking on the paradox of asking a masked man who he is.

Evey: Oh, right...

V: But on this most auspicious of nights, permit me then, in lieu of the more commonplace sobriquet, to suggest the character of this *dramatis persona*: Voilà! In view, a humble vaudevillian veteran [...] it is my very good honour to meet you and you may call me V.

Evey: Are you like a crazy person?

V: I am quite sure they will say so.

V for Vendetta (the film)

II Non-prototypical combinations of values

1 Introduction

As discussed in Sections I.4.2.3.1-2, the two parameters that underlie the conceptual space for 'who?' and 'what?', viz. ENTITY TYPE and TYPE OF REFERENCE (& EXPECTED ANSWER), can give rise to non-prototypical combinations of values with respect to the choice of a non-selective interrogative pronominal (henceforth, *non-prototypical combinations of values*). Instead of the prototypical combinations of values, questions can involve the following, non-prototypical combinations: [person + classification (+ common noun)] and [thing + identification (+ proper name)]. In some languages, the value [person] is extended to include animate things, giving rise to a third, minor instance of a non-prototypical combination of values, viz. [animate thing + classification (+

common noun)]. Figure 1 gives a simplified overview, leaving out animate things. The greyed cells are those in which a non-prototypical combination of values arises.

Figure 1.

		ENTITY TYPE	
		[person]	[thing]
TYPE OF REFERENCE (& EXPECTED ANSWER)	[identification (+ proper name)]	WHO?	
	[classification (+ common noun)]		WHAT?

There are three possible strategies to resolve these non-prototypical combinations of values. Either the choice between ‘who?’ and ‘what?’ is avoided (see Section I.4.2.3.1), or ‘who?’ dominates, or ‘what?’ dominates. This chapter provides a cross-linguistic examination of the latter two strategies. Section II.2 discusses the choice for ‘what?’ in the non-prototypical value combination [person + classification (+ common noun)] and Section II.3 treats the choice for ‘who?’ in questions asking for [thing + identification (+ proper name)] (schematized in Figure 2). That is, these two sections treat the situations in which the parameter TYPE OF REFERENCE is decisive for the choice of an interrogative.

Figure 2.

		ENTITY TYPE	
		[person]	[thing]
TYPE OF REFERENCE (& EXPECTED ANSWER)	[identification (+ proper name)]	WHO?	who? (Section II.3)
	[classification (+ common noun)]	what? (Section II.2)	WHAT?

The use of ‘who?’ for the non-prototypical value combination [thing +

identification (+ proper name)] is illustrated in (1) from Kgalagadi, and the use of ‘what?’ for [person + classification (+ common noun)] in (2) from English.

Kgalagadi (Niger-Congo, Bantu S30; Botswana; Kems Monaka, p.c.)

- (1) [A:] *libizho la lehelo lo ke anye?*
G5.name AG5.CON G11.place AG11.DEM COP who

[B:] *ke Hukhuntsi*
COP PROP

‘[A:] What (lit.: ‘who?’) is the name of this place? [B:] It’s Hukhuntsi (a village name)’

English

- (2) *What is he actually? A doctor?*

The only situation in which the parameter ENTITY TYPE is decisive that will be discussed here (in Section II.4) is that in which the value [person] is extended to include animate things, as manifested through the choice for ‘who?’, giving rise to a third, minor instance of a non-prototypical combination of values, viz. [animate thing + classification (+ common noun)], as shown in Figure 3.

Figure 3.

		ENTITY TYPE		
		[person]	[animate thing]	[inanimate thing]
TYPE OF REFERENCE (& EXPECTED ANSWER)	[identification (+ proper name)]	WHO?		
	[classification (+ common noun)]		who? (Section II.4)	WHAT?

The situation schematized in Figure 3 can be illustrated with example (3) from Russian.



Russian

- (3) *Kto eto tebja ukusil?*
who this you bit

‘[Looking at a swelling on someone’s hand clearly caused by an insect bite:] What stung you? (e.g., a wasp, a bee, etc.)’

Other situations in which the parameter entity type is decisive, as schematized in Figure 4, will not be discussed explicitly. This is because, normally, descriptions appear to take it for granted that in language L, a given interrogative X glossed as ‘who?’ is used only for persons and a given interrogative Y glossed as ‘what?’ is used only for things, thus implying, as it were, that ENTITY TYPE is the only parameter governing the choice between ‘who?’ and ‘what?’.¹ Therefore, I believe it will be more enlightening to take a different perspective, which regularly remains unconsidered.

Figure 4.

		ENTITY TYPE	
		[person]	[thing]
TYPE OF REFERENCE (& EXPECTED ANSWER)	[identification (+ proper name)]	WHO? 	what? (not discussed)
	[classification (+ common noun)]	who? (not discussed)	 WHAT?

The use of ‘who?’ for the non-prototypical value combination [person + classification (+ common noun)] is illustrated in (4) from Russian, and the use of ‘what?’ for [thing + identification (+ proper name)] in (5) from English.

Russian

- (4) *A on kto voobščē? Doktor?*
 and he who actually doctor
 ‘What is he actually? A doctor?’

English

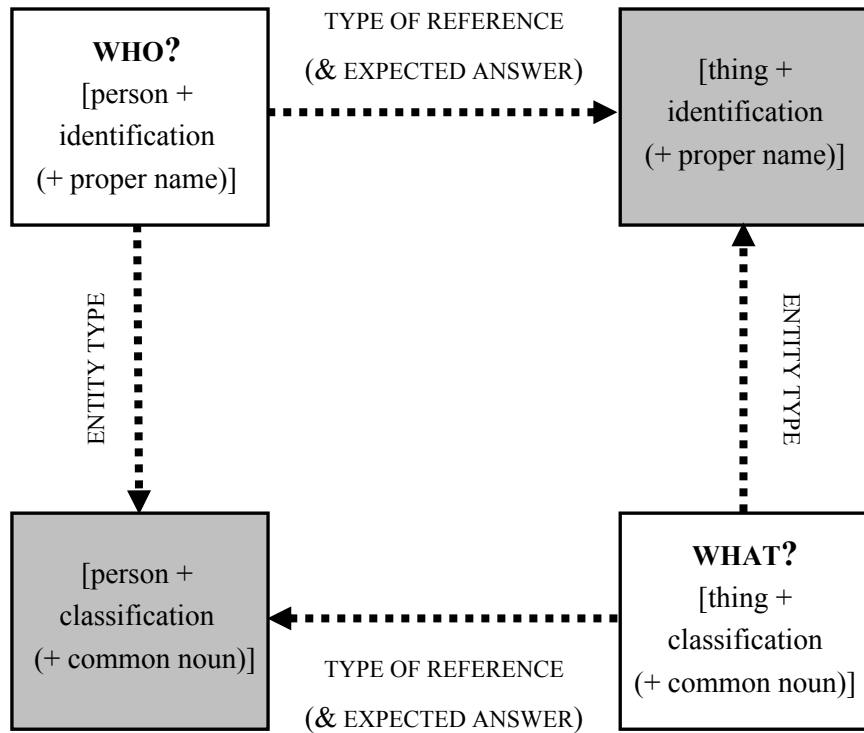
- (5) *What is the name of this place? What is this place called?*

Figures 1-2 and 4 can be summarily represented in a slightly different form, as in Figure 5. Here, the prototypical combinations of values associated with ‘who?’ and ‘what?’ are found in the white boxes and the non-prototypical combinations of values in the greyed boxes. The white boxes are linked to the

¹ To be more precise, though, it is not so much that descriptions take something for granted but rather they simply do not consider the possibility of any variation between different languages in the domain of non-selective interrogative pronominals.

greyed ones by dominance relations represented by the dashed arrows. The parameter next to the arrow is the parameter governing the choice between ‘who?’ and ‘what?’ in a given case.

Figure 5. Conceptual space for non-selective interrogative pronominals (in terms of prototypical and non-prototypical combinations of values and ‘who?’/‘what?’-dominance)



Section II.5 provides a conclusion and a typology of languages according to their resolution strategies. For convenience sake, I will sometimes refer to the three non-prototypical combinations of values with the following shorthands (6).

- (6) a. KIND-questions → [PERSON + CLASSIFICATION (+ COMMON NOUN)]
 b. NAME-questions → [THING + IDENTIFICATION (+ PROPER NAME)]
 c. ANIMATE-questions → [ANIMATE THING + CLASSIFICATION (+ COMMON NOUN)] (& ‘who?’-dominance)

Before proceeding to the discussion, a few general remarks need to be made. Thus, although on the whole, the possible kinds of non-prototypical combinations of values are determined by the way the conceptual space for ‘who?’ and ‘what?’ is organized, in the languages of the world more variation in the use of ‘who?’ and ‘what?’ is observed than could have been presumed based only on the possible kinds of non-prototypical combinations of values. The reason is that the same non-prototypical combination of values may arise in

different contexts and languages may prefer to resolve such a non-prototypical combination of values through ‘who?’-dominance in one context but ‘what?’-dominance in another (not to mention various avoidance strategies that may be possible in any context), even though in principle, the kind of non-prototypical combination of values in both contexts is the same. Thus, a non-prototypical combination of values of the kind [person + classification (+ common noun)] may arise when the person at issue is asked to be classified according to his/her occupation, sex, ethnic origin, social position, relation to another person (e.g., a kinship relation), etc. Some languages may be consistent in preferring in all such contexts either ‘who?’, as for instance Russian does, or ‘what?’, as for instance English does, whereas some languages may prefer ‘who?’ in some contexts but ‘what?’ in other contexts (avoidance strategies left aside). Vietnamese, for instance, in the case of a non-prototypical combination of values of the kind [person + classification (+ common noun)], selects *ai* ‘who?’ (or an avoidance strategy) in all contexts but one, viz. a question on the (kinship) relation of the person at issue to another person, where only *gì* ‘what?’ is possible, cf. (7).

Vietnamese (Austro-Asiatic, Mon-Khmer, Viet-Muong; Vietnam; Thu Thi Anh Nguyen, p.c.)

- (7) [A:] *Mary là gì/*ai của bạn?*
 Mary is what/*who of you
 [B:] *Chị ấy là chị dâu của tôi*
 she is sister-in-law of me
 ‘[A:] What is Mary to you? [B:] She is my sister-in-law.’

Admittedly, given that my data on non-prototypical combinations of values have been mostly gathered by means of a questionnaire, I may have missed some relevant contexts. Similarly, given that I have been able to consult only a very restricted number of speakers for each language (usually just one), there is a danger that for some languages my data may eventually prove to be somewhat skewed in being more restrictive than the actual language use is in reality. Some speakers may reject a given interrogative utterance not because of its lower degree of grammaticality but because they have stronger ideas (which are not necessarily correct) about what is the most idiomatic way of asking this or that question, so that they may tend to equate what they perceive as a lower degree of idiomaticity with a lack of grammaticality.² Admittedly, the border between the

² For instance, several speakers of German I consulted rejected the possibility of literally translating *What is Mary to you?* (as a question inquiring about the kinship relation between the interlocutor and Mary, e.g. *She is my sister-in-law*) as *Was ist Mary für dich?*. Instead, they usually preferred something like *Wie ist Mary mit dir verwandt?* ‘How is Mary related to you?’ (lit.: ‘how is Mary with you related’). However, other speakers saw no problem in using

two is not always easy to draw. Furthermore, besides idiomaticity, an important role may be played by various cultural restrictions. Certain kinds of questions may be characterized as ungrammatical because asking certain things is “not done” or because the relevant realia are not applicable in a given linguistic community. For instance, in many societies, it is considered impolite to ask for someone’s name, at least directly. In Ahaggar Tuareg (Afro-Asiatic, Berber; Algeria), it is impolite to ask about the sex of a newborn, “you must assume it is a boy unless you are told otherwise” (Karl-G. Prasse, p.c.). In many traditional societies, people do not have jobs, in the sense of regular remunerative positions, so that one will never be asked to classify a given person according to his or her occupation.

I have tried to reduce the chances for missing relevant contexts and to allow for a cross-check of “excessively” restrictive judgements by including in the questionnaire some very general questions such as (8) and (9).

- (8) *Does an equational construction of the English type ‘What is John?’ (i.e., where the interrogative pro-noun ‘what?’ is equated to a noun designating a person) make any sense at all in your language?*
- (9) *Does an equational construction of the type ‘Who is X?’ (where X is a noun that does not designate a person, by preference it designates a thing) make any sense at all in your language?*

Sometimes, the validity of a given judgement in a questionnaire on language L may be further supported in a less direct way. Thus, the Vietnamese case mentioned above may be further corroborated by the fact that a rather similar situation appears to exist in Mandarin Chinese (Sino-Tibetan, Chinese; Renzhi Li, p.c.), Taiwanese Min Nan Chinese (Sino-Tibetan, Chinese; Cheng-Fu Chen, p.c.) and Japanese (Mizuki Miyashita, p.c.; Nobuko Yoneda, p.c.).

Besides the questionnaire, I have also consulted numerous descriptions of particular languages. However, as I have already pointed out elsewhere, most sources provide very little explicit or implicit (in the form of examples)

Was ist Mary für dich? In this respect, consider also example (i).

German

- (i) [*Und, lieber Jürgen, du weißt ja, dass ich deine Beccibilder sehr schätze.*]

Was ist sie für dich, deine Nichte, Schwester, Tochter?

what is she for you your niece sister daughter

‘[(A reaction on a series of pictures of a girl called Becci posted on a website) And, dear Jürgen, well, you know that I very much appreciate your Becci pictures.] What is she to you, your niece, sister, daughter?’ (<http://www.fotocommunity.de/pc/account/myprofile/532495>, retrieved 28.04.2007)

information on the semantics and patterns of use of the interrogative pronominals. Finally, although my sample is probably sufficient to discern the major types of relevant contexts and the general patterns of areal and genealogical distribution, some interesting details may still be missing.

2 ‘What?’-dominance in KIND-questions

This section examines KIND-questions, i.e. questions involving a non-prototypical combination of values of the kind [person + classification (+ common noun)], and languages that at least in some contexts, prefer to use ‘what?’ rather than ‘who?’ in KIND-questions.

I will begin this section by discussing in Section II.2.1 four common types of contexts where this non-prototypical combination of values arises: questions about someone’s (i) biological affiliation (i.e., sex), (ii) social affiliation (e.g., ethnic origin, clan, caste, denomination, etc.), (iii) functional affiliation (e.g., occupation, role, etc.), and (iv) personal affiliation (i.e., someone’s relation to another person, especially of a kin kind). In Section II.2.2, I will present some common preferences and particular restrictions related to the form of the constructions used in these contexts as regards the use of ‘what?’. Finally, in Section II.2.3 a brief summary will be provided.

2.1 Contexts

2.1.1 Biological affiliation: sex

One of the contexts where the non-prototypical combination of values of the kind [person + classification (+ common noun)] arises is represented by questions about someone’s sex. Quite a few languages appear to prefer ‘what?’ to ‘who?’ in this situation, as illustrated in (1) from Ma’di and (2) from Nez Perce.

Ma’di (Nilo-Saharan, Central Sudanic, East; Uganda; Mairi J. Blackings, p.c.)

- (1) a. *bārā r̄ àdũ? zángwá dʒɔ bāragó?*
 child DEF what girl or boy
 ‘What is the baby? A boy or a girl?’
 b. *màdĩ dʒɔ gá nà àdũ? ízí dʒɔ ágó?*
 person house LOC that what woman or man
 ‘What is the person in that house? A man or a woman?’

Nez Perce (Plateau Penutian, Sahaptin; USA; Noel Rude, p.c.)

- (2) a. *?itúu ?ewc’éeye miyapkáawit? háacwal ?ít’qo pit’ín’?*
 what her.baby is.born boy or girl
 ‘What did her baby become? A boy or a girl?’
 b. *kaa ?itúu híiwes kíi? háama ?ít’qo ?áayat?*
 and what is this man or woman
 ‘And what is this one? A man or a woman?’

In Ma'di and Nez Perce, 'what?' is used in both questions about the sex of a baby and questions about the sex of an adult. However, in some languages only the former use appears to be possible, as illustrated in (3) from Kwaza and (4) from Malayalam.

Kwaza (Isolate; Brazil; Hein van der Voort, p.c.)

- (3) a. *tsũhũ-ratĩ-re tswá-tohói-re etáy-tohói-re?*
 what-FOC-Q man-CLF.SMALL-Q woman-CLF.SMALL-Q
 'What is it? Is it a boy or a girl?'
 b. (**tsũhũ-ratĩ-re*) *tswá-re etáy-re?*
 (*what-FOC-Q) man-Q woman-Q
 'Is it a man or a woman?'

Malayalam (Southern Dravidian; India; M. T. Hany Babu & Menon Mythili, p.c.)

- (4) a. *kuṭṭi eṇṭ-aṇə? aṇ-oo peṇṇ-oo?*
 child what-is male-or female-or
 'What is the baby? A boy or a girl?'
 b. [A: *ñaan innale paṭṭaalatt-il chern-u. Pakuthi officer-maarə aṇuṇṇaḷ aṇə baki peṇṇuṇṇaḷum.* B:]
commandar aarə aṇə? [A:] aṇ-aṇə
 commander who is male-is
 '[A: I joined the army yesterday. Half of the officers are men, the others are women. B:] And what is the commander? [A:] It's a man.'

What is more, in some languages, questions about the sex of a child appear to represent the only context where 'what?' can be applied to a person, as for instance, in Ikwere (5) and Zo'é (6).

Ikwere (Niger-Congo, Benue-Congo, Igboide; Nigeria; Sylvester Osu, p.c.)

- (5) *ḡ m̀r̀r̀ ḱńí? r̀̀k̀á sà r̀̀ỳá?*
 3SG engender-PST what male or female
 'What did she give birth to? A boy or a girl?'

Zo'é (Tupí, Tupí-Guaraní, Subgroup VIII; Brazil; Ana Suelly Arruda Câmara Cabral, p.c.)

- (6) *mo?é dugét? kuña kubi?é?*
 what is.like female male
 'What is it? (lit.: 'What is it like?') A girl or a boy?'

It is not uncommon cross-linguistically for children to be treated as lower in animacy than adults. In all probability, the difference between children and adults as regards the use of 'who?' and 'what?' in questions about sex is another

manifestation of this tendency.³

2.1.2 *Social affiliation*

Questions on the social affiliation of a person represent another common kind of contexts where the non-prototypical combination of values of the kind [person + classification (+ common noun)] arises. The social affiliation of a person may be based on his or her ethnic origin, clan, caste, denomination, etc. Consider, for instance, (7-9) from Mayrinax Atayal and Wichita, where ‘what?’ rather than ‘who?’ is used to ask about the person’s ethnic origin, and (9) from Poligus Evenki, where ‘what?’ is used to ask for the person’s classification in terms of clan, family group.

Mayrinax Atayal (Austronesian, Atayalic; Taiwan; Huang 1996:273)

- (7) [A:] *nanuwan=su?* [B:] *itaal=cu*
 what=2SG.NOM Atayal=1SG.NOM
 ‘[A:] What tribeship are you? (lit.: ‘What are you?’) [B:] I am Atayal’

Wichita (Northern Caddoan; USA; David S. Rood, p.c.)

- (8) *e:kinni:ʔi?*
 |*e:kiri-na-i:-ʔi*|
 what-PRS.Q-INDF.SBJ-be
 ‘[A:] What is he/she? [B: He/she is Wichita]’

Poligus Evenki (Northern Tungusic; Russia; Konstantinova 1968:73)

- (9) *e:kun bi-si-nni? bi Kurkogir bi-si-m*
 what be-PRS-2SG 1SG PROP be-PRS-1SG
 ‘[A:] What are you? [B:] I am a Kurkogir (i.e., I belong to the Kurkogir clan, the name of my family is Kurkogir)’

Note that in Wichita, for instance, it appears to be possible to apply ‘what?’ to a person only in the context of a question about the person’s “tribal affiliation” or the sex of a child. However, I found no languages where the use of ‘what?’ for persons is confined to questions on social affiliation.

2.1.3 *Functional affiliation*

The next typical context where the non-prototypical combination of values of the

³ It may also be worth pointing out, in this respect, that whereas people may make mistakes in telling the sex of (clothed) children, especially babies, such mistakes are much less common with respect to adults. In this connection, a question about the sex of a child would to a certain extent be more natural than a question about the sex of an adult.

kind [person + classification (+ common noun)] arises is represented by questions on the “functional” affiliation of a person. First of all, questions on functional affiliation include questions on the person’s occupation, job, as in (10) from Dutch, or position, as in (11) from Nganasan.

Dutch (Haeseryn *et al.* 1997:§5.7.4.2)

- (10) *met wat is ze nu getrouwd? een arts of een apotheker?*
 with what is she now married a doctor or a druggist
 ‘With what is she married now? A doctor or a druggist?’

Avam Nganasan (Uralic, Samoyedic; Russia; Valentin Goussev, p.c.)

- (11) [A:] *takəə-tə bəjkaʔa maa?* [B:] *setəgəə*
 DEM-2SG.GEN husband what director
 ‘[A:] What is her husband? [B:] A director’⁴

Furthermore, I also subsume under questions on functional affiliation of a person questions asking for a classification of a person according to various other features that do not directly concern the person’s social affiliation or the person’s relation to another particular person and that do not serve to express the persons’ sex, even if as such, they may be applicable only to persons of one sex or a given social group. In this respect, consider, for instance, examples (12) and (13).

Classical Arabic (West Semitic, Central; Brockelmann 1913:195)

- (12) *fa-ma: tazawwaj-ta bikr-a-n ʔaw*
 so-what marry.PRF-2SG.M virgin-ACC.SG-INDEF or
tayyib-a-n?
 deflowered-ACC.SG-INDEF
 ‘What have you married, a virgin or an already deflowered one?’

Shehri (West Semitic, South, Eastern; Oman; Bittner 1917:106-107)

- (13) *iné tīt-k?*
 what wife-2SG.M
 ‘[And he asked him,] What (kind of person) is your wife? [He said to him, ‘She is worth nothing, she has damaged my clothes.’]’

I found no languages where the use of ‘what?’ for persons is confined to questions on functional affiliation.

2.1.4 Personal affiliation

The last typical context where the non-prototypical combination of values of the

⁴ The suffix *-tə* 2SG.GEN functions as a kind of definite article in Nganasan.

kind [person + classification (+ common noun)] arises is represented by questions on the “personal” affiliation of a person, i.e. the person’s classification according to the person’s relation to another person. Relation through kinship clearly represents the most typical case of such an interpersonal relation. In this respect, consider (14) from Bamana and (15) from Kayabí.

Bamana (Niger-Congo, Western Mande; Mali; Dumestre 2003:151)

- (14) [A:] *í mún dòn?* [B:] *ń dógóké` dén` dòn*
 2SG what here.is 1SG younger.brother-ART child-ART here.is
 ‘[A:] What is he to you? (lit.: ‘Here is your what?’) [B:] It’s my younger brother’s child’

Kayabí women’s speech (Tupí, Tupí-Guaraní, Subgroup V; Brazil; Rose Dobson, p.c.)

- (15) [A:] *ma’jate Maria kyn a ene upe?*
 what Q PROP 3SG.F 2SG to
 [B:] *je-rewirera kĩa remirekoa futat kyn a*
 1SG-brother 3SG.M wife EMPH 3SG.F
 ‘[A:] What is Mary to you? [B:] She is my sister-in-law’

Note that questions such as French (16a) do not ask about the person’s personal affiliation but rather about the person’s functional affiliation, since the person in question is clearly unrelated to the interlocutor.⁵ Example (16a) can be compared to (16b), where the speaker suspects that there is some relation between the interlocutor and the person in question.

French (based on Isabelle Bril & Christiane Pilot-Raichoor, p.c.)

- (16) a. *Qu’est-ce que Staline pour toi?*
 what-is-DEM REL.ACC PROP for you
 ‘What is Stalin to you?’ (the question is about Stalin’s role in history, the interlocutor’s appreciation of Stalin’s personality, etc.)
 b. *Qu’est-ce que Marie pour toi? une copine? ta maîtresse?*
 what-is-DEM REL.ACC PROP for you a friend your mistress
 ‘[A jealous woman:] What is Mary to you? a friend? your mistress?’

In some languages, questions about personal affiliation appear to represent the only context where ‘what?’ may normally be applied to a person. Thus, as was

⁵ Alternatively, the personal proper name *Staline* may be interpreted to be used as an autonym in (16a), i.e. as ‘the name/word *Staline*’, so that the question would be ‘What does *Staline* mean (according) to you?’ (cf. the discussion of (52) in Section I.4.2.3.3.3).

mentioned in Section II.1, this seems to be the case in Vietnamese (Austro-Asiatic, Mon-Khmer, Viet-Muong), Mandarin Chinese (Sino-Tibetan, Chinese), Taiwanese Min Nan Chinese (Sino-Tibetan, Chinese) and Japanese. Consider also (17) from Sateré-Mawé, where the situation seems to be the same.

Sateré-Mawé (Tupí; Brazil; Sérgio Meira, p.c.)

- (17) [A:] *kan ee-pe mejūwa, pi'ā hiin, Maria?*
 what.Q 2-to DEM girl little PROP

[B:] *Maria uhe-pe ui-mēpít*
 PROP 1-to 1-F.daughter

‘[Person A:] What is this one to you, the little girl Maria? [Woman B:] Maria to me, she is my daughter’

2.2 Constructions

Languages may differ as to the kinds of constructions in which they prefer or allow KIND-questions to be formulated with ‘what?’ in a particular context. For instance, it appears that several languages which use ‘what?’ in questions about the sex of a child (but not an adult) do not allow for ‘what?’ to be used as a nominal predicate in a construction of the kind ‘What is the baby?’. Instead, something like ‘What did she give birth to?’ must be resorted to, as in (5) from Ikwere (Niger-Congo, Benue-Congo, Igboide; Nigeria) above or in (18) from Georgian here.

Georgian (Kartvelian; Georgia; Nino Amiridze, p.c.)

- (18) (*mas*) *ra* *eqola,* *gogo* *tu bič-i?*
 (she.DAT) what.NOM (s)he.it.got.born.to.him/her girl.NOM or boy-NOM
 ‘What did she give birth to? A girl or a boy?’

No such restriction has been found in other contexts. Quite to the contrary, my impression is that in other contexts the predicative use of ‘what?’ is more typical than the non-predicative one (for the latter use see, e.g., (12) above). Note, however, that in questions about the social or functional affiliation of a person, some languages appear to disallow a proper name (and sometimes also a common nominal) as the subject of a predicatively used ‘what?’. No such restriction appears to exist for questions on the personal affiliation of a person. In Badaga, for instance, (19) with a proper name *Bīma* as the subject and *ēna* ‘what?’ as the nominal predicate is basically a ‘what’s up?’-construction (cf. Section I.4.2.3.3.3) and not a question about *Bīma*’s classification.

Badaga (Southern Dravidian; India; Christiane Pilot-Raichoor, p.c.)

- (19) [A:] *Bīma ēna?* [B:] *Bīma buddu buṭṭa*
 PROP what PROP fall.PTCP AUX.PFV.PST.3SG
 ‘[A:] What’s up with Bima? (What happened to Bima?) [B:] Bima has fallen’

At the same time, a seemingly identical structure with a deictic subject, as in (20a) and (21a), can be used to ask about the sex of a child and the functional affiliation of a person respectively. For the sake of comparison, I also provide a question on the sex of an adult with a common nominal as the subject (20b), a question on the functional affiliation of a person with a proprial subject (21b), and a question on the personal affiliation of a person with a proprial subject (22).

Badaga (Christiane Pilot-Raichoor, p.c.)

- (20) a. *idu ēna, eṇṇ-ā, gaṇḍ-ā?*
 DEM.NON<HUM> what girl-Q boy-Q
 ‘What is it, a boy or a girl?’
 b. [A: *ninne nā militari sēde, alliyō pādiya emmukaru pādiya gaṇḍumukaru*]
 [B:] *alli commandar dāra-v-ā?* [A:] *adu manca*
 there what who-AFF-Q DEM.NON<HUM> man
 ‘[A: I joined the army yesterday. Half of the officers are men, the others are women. B:] And what is the commander? [A:] It’s a man’
- (21) a. [A:] *ama ēna?* [B:] *klarku*
 DEM.M.SG what clerk
 ‘[A:] What is he? [B:] A clerk’
 b. *Bīma dāra? tacca-n-ā?*
 PROP who carpenter-AFF-Q
 ‘What is Bima? A carpenter?’
- (22) [A:] *Mādi ninaga ēna āpadu?* [B:] *enna attige*
 PROP 2SG.DAT what become.NMLZ 1SG.OBL sister-in-law
 ‘[A:] What is Mary to you? [B:] She is my sister-in-law’

Consider also examples (23-25) from Muna, which is similar to Badaga as regards the restriction on the use of a proprial subject with a predicative ‘what?’ in all contexts except questions on personal affiliation. However, Muna differs from Badaga in allowing a predicative ‘what?’ with common subjects in questions on the person’s functional affiliation (23b) and the sex of an adult (24). The Muna example (25) is a question on the personal affiliation of a person with a proprial subject and can be compared to the Badaga example (22).

Muna (Austronesian, Malayo-Polynesian, Sulawesi; Indonesia; René van den Berg, p.c.)

- (23) a. **o hae la Amir maitu?*
 ART what ART PROP DEM
 ‘What is Amir?’
- b. [A:] *o hae ko-toko-ghoo-no maitu?*
 ART what have-shop-APPL-ACT.PTCP DEM
 [B:] *ingka o polisi*
 as.you.know ART police
 ‘[A:] What is the shop owner? [B:] As you know, he is a policeman’
- (24) [B:] *neangka-mu o hae?* [A:] *o moghane*
 leader-2SG ART what ART man
 ‘[A: I joined the army yesterday. Half of the officers are men, the others are women. B:] And what is your commander? [A:] A man’
- (25) [A:] *Hae-no-mu wa Maria maitu?* [B:] *isa-ku*
 what-3SG-2SG ART PROP DEM older.sibling-1SG
 ‘[A:] What is Maria to you? [B:] She is my older sister’

The restriction in some languages on the use of proprial and sometimes also common personal subjects with a predicative ‘what?’ in all contexts except questions on personal affiliation should probably be viewed as another manifestation of the correlations already discussed in Sections I.2.5-2.6. Thus, there is a semantic link between the category of proper names, the concept of personhood and ‘who?’, on the one hand, and the category of (non-personal) common nouns, the concept of thingness and ‘what?’, on the other. At the same time, there is no such link between the category of deictics and either the concept of personhood or the concept of thingness. In other words, whereas some languages will never allow ‘what?’ in questions about persons, irrespective of the way the person is referred to in the question, some languages may go a bit further in tolerating ‘what?’ in questions about persons only when the person is referred to with a deictic, because with the latter the clash between the personal reference and the non-human semantics of ‘what?’ is the smallest, as compared to personal common nouns and especially personal proper names.

The exclusion from the tendency just discussed of questions on personal affiliation, as illustrated by (22) and (25), is probably due to the fact that the prototypical answer to such questions is a kinship term and kinship terms are functionally rather close to deictics. Just like a deictic locates an entity with respect to some deictic centre within a given space, a kinship term locates a person with respect to another person within a given kinship network.⁶

⁶ Note in this respect that in some languages kinship terms may function as personal pronouns.

2.3 ‘What?’-dominance in KIND-questions: summary

In Section II.2.1, I have distinguished four common types of contexts where KIND-questions are formulated ‘what?’ in the languages of my sample: (i) questions about someone’s biological affiliation (i.e., sex), (ii) social affiliation (e.g., ethnic origin, clan, caste, denomination, etc.), (iii) functional affiliation (e.g., occupation, role, etc.), and (iv) personal affiliation (i.e., someone’s relation to another person, especially of a kin kind). As discussed in Section II.2.2, from a cross-linguistic perspective, there appears to be only one interesting restriction related to the form of the constructions used in KIND-questions formulated ‘what?’, viz. the restriction in some languages on the use of proprial and sometimes also common personal subjects with a predicative ‘what?’ in all contexts except questions on personal affiliation.

As discussed in Section II.2.1, some languages allow (or require) KIND-questions formulated with ‘what?’ only in some of the four types of possible contexts. It appears possible to organize most of these contexts in an implicational hierarchy as regards the possibility of ‘what?’-dominance in KIND-questions (26).

(26) *‘what?’-dominance hierarchy for KIND-question contexts:*

sex of an infant < social affiliation < functional affiliation < sex of an adult

Note that this hierarchy represents only a (strong) tendency rather than an absolute universal. Furthermore, it does not take into account avoidance strategies. In other words, this hierarchy implies that if, for instance, a given language allows for KIND-questions to be formulated ‘what?’ when the question is about the functional affiliation of a person, in all the contexts to the left it will also either allow for ‘what?’ or will use an avoidance strategy, but it will not require only the use of ‘who?’.

As can be observed, whereas questions on the sex of an infant are on the left end of the hierarchy in (26), questions on the sex of an adult are on its right end. There appear to be two reasons for this. First, infants are often treated differently from adults as regards animacy (cf. Section II.2.1.1). Second, questions on the sex of an adult have often been perceived as rather unnatural by respondents of the questionnaire, which to a large extent, may account for the frequent use of the regular human interrogative ‘who?’ here even when in all other contexts ‘what?’-dominance was clearly preferred in KIND-questions.

Another detail of the hierarchy in (26) that may strike the eye is the absence of context “personal affiliation”. It has proved difficult to place this context

In Vietnamese, for instance, *anh* ‘older brother’ and *chị* ‘older sister’ can be used as the second or third person singular pronoun in reference to a young man and young woman respectively.

somewhere on the hierarchy without running the risk of reducing too much the strength of the implicational relations. In all probability, the reason for this somewhat deviant behaviour of personal affiliation is that it involves classification in terms of *relation* between two persons, whereas other kinds of contexts involve only classification of a person as such. This distinction can be compared to the distinction between transitive and intransitive verbs or between inalienable and alienable possession.

By way of conclusion, let us consider the areal and genetic distribution of the languages in my sample that at least in some contexts, allow (or require) ‘what?’-dominance in cases of the non-prototypical combination of values [person + classification (+ common noun)]. Languages with ‘what?’-dominance in KIND-questions appear to be common in Africa. Another area can be conveniently described as something like the “southern belt” of Eurasia. This “belt” largely follows the traditional distribution of the Indo-European languages (with a noticeable exception of many Slavic languages) and continues further to East and Southeast Asia where it can be further extended into the area covered by Austronesian languages. In northern Eurasia, there appears to be only one major island of languages using ‘what?’ in KIND-questions, viz. Tungusic languages. Languages with ‘what?’-dominance in KIND-questions do not seem to be common in Australia and New Guinea. The same seems to be the case in the Americas, although the situation is less clear there. The overall picture in the Americas may be somewhat disturbed by the fact that English and the Romance Indo-European languages strongly dominant in the region today all use ‘what?’ in KIND-questions.

3 ‘Who?’-dominance in NAME-questions

3.1 Preliminaries

This section examines NAME-questions, i.e. questions about proper names (and identification) of things which involve a non-prototypical combination of values of the kind [thing + identification + proper name]. Especially, I am interested here in languages that at least in some contexts, prefer to use ‘who?’ rather than ‘what?’ in NAME-questions, as in (1) from Namia.

Namia (Sepik-Ramu, Sepik, Yellow River; Papua New Guinea)

- (1) a. [A:] *ne-k(a) ilei tal(a)?* [B:] *John*
 2SG-POSS name who PROP
 ‘[A:] What is your name? [B:] John’ (Becky Feldpausch, p.c.)
- b. *tala ne-m naked-e?*
 who 2SG-OBJ tell-NON<FUT>
 ‘Who told you?’ (Feldpausch & Feldpausch 1992:73)

The interrogative *tala* ‘who?’ in (1) questions the identity of *ne-k ilei* ‘your name’. That is, (1) involves a non-prototypical combination of values of the kind [thing + identification + proper name]. In this respect, note that the proprial nominal *John* in the answer is used autonomously, i.e. it does not refer to anything but itself, viz. ‘the name *John*’. That is, the autonym *John* is, as it were, the proper name of the proprial nominal *John*.⁷

As can be readily observed from the translation of (1), English opts for *what*, since the entity questioned about is a thing and not a person. The use of *tala* ‘who?’ in the Namia original can be explained by the fact that the entity questioned about is a *personal* proper name. That is, in this context, Namia takes into account the so-called “categorical presuppositional meaning” of proper names (cf. Van Langendonck 2007:79-81), i.e. the category of entities which the use of the particular proper name presupposes.⁸

That contexts like (1) involve a non-prototypical combination of values is further supported by the fact that there are many languages that will allow neither ‘what?’ nor ‘who?’ here. Instead, an avoidance strategy will be used. Given that

⁷ According to Van Langendonck (2007:246-249), autonyms behave in many respects like proper names and should be considered as “a subclass of proper names in their own right” (2007:95).

⁸ Here are a few more examples. The categorical presuppositional meaning of the proper name *Paris*, for instance, would be ‘city’, as becomes apparent in *the City of Paris*, whereas the categorical presuppositional meaning of the proper name *Fido* would be ‘dog’, as can be seen in the appositional structure *Fido the dog* (Van Langendonck 2007:80-81).

‘name’ is a thing said, the great majority of languages preferring an avoidance strategy appear to opt for an interrogative proadverbial ‘how?’, as in (2) from Arabela. The use of ‘how?’ seems to suggest that something like ‘How is your name said/pronounced?’ is implied here.

Arabela (Zaparoan; Peru; Rolland Rich, p.c.)

- (2) [A:] *taa-te quia sesa-ni?* [B:] *John*
 how-Q 2SG name-Q PROP
 ‘[A:] What is your name? (lit.: ‘How is your name?’) [B:] John’

Other common possibilities include the use of an interrogative proadverbial ‘where?’, as in (3) from Hausa,⁹ or a selective interrogative pronominal ‘which one?’, as in (4) from Amharic.¹⁰

Standard (Eastern) Hausa (Afro-Asiatic, West Chadic; Nigeria)

- (3) *inaa suuna-n-ka?*
 where name-of-2SG
 ‘What is your name? (lit.: ‘Where is your name?’)’ (Paul Newman, p.c.)

Amharic (Afro-Asiatic, West Semitic; Ethiopia; Seyoum Mulugeta, p.c.)

- (4) *səm-əh yätəṇṇaw nəw?*
 name-2SG which.one.M.SG COP.M.SG
 ‘What is your name? (lit.: ‘Which one is your name?’)’

I believe the use of ‘where?’ and ‘which one?’ should be explained through the notion of selection. To begin with, note that cross-linguistically, there is a recurrent link between ‘where?’ and ‘which one?’, since both interrogatives expect a deictic expression as an answer, viz. either an adverbial, such as ‘there’, or a pronominal, such as ‘that (one)’.¹¹ A selective interrogative pronominal or a semantically related locative interrogative proadverbial appears to be used in questions about proper names to avoid the non-prototypical combination of

⁹ It is also possible to ask the same question using *yàaya* (Paul Newman, p.c.), and at least in Central Hausa (Katsinanci dialect; Niger), one can use *mì(i)* ‘what?’, as in *mì-nee nèe suuna-n-kà* ‘What is your name?’ (lit.: what-M.SG COP.M.SG name-of-2SG.M) (Mahamane Laoualy Abdoulaye, p.c.).

¹⁰ It is also possible, and perhaps, even more common to ask the same question as *səm-əh man nəw* (name-2SG who COP.M.SG) ‘What is your name?’ (lit.: ‘Who is your name?’) or *man tə-balal-əh* (who 2SG.M-called-2SG.M) ‘What are you called?’ (lit.: ‘Who are you called?’) (Seyoum Mulugeta, p.c.; Joachim Crass, p.c.).

¹¹ Compare, for instance, [There are ten apples lying on the table:] *Where is the apple you would like?* and *Which (one) is the apple you would like?* (or *Which apple you would like?*). The semantic difference between *where* and *which (one)* appears to be minimal here.

values of the kind [thing + identification + proper name] by replacing [identification] with [selection].

In what follows, Section II.3.2 will discuss the constructions in which the non-prototypical combination of values [thing + identification + proper name] may arise. In Section II.3.3, I will present the kinds of contexts (i.e., the kinds of proper names) for which ‘who?’-dominance in this non-prototypical combination of values has been found. Finally, in Section II.3.4, concluding remarks will be made.

3.2 Constructions

The construction illustrated in (1) above can be conveniently called [*IPW is X’s name?*]-construction. Besides this construction, some other constructions also may involve the same kind of non-prototypical combination of values, viz. (i) [*IPW is X? (by name?)*], where *X* is a thing, as in (5a), and (ii) [*IPW is X named/called?*] (or [*IPW do they name/call X?*], [*IPW does X name/call himself/herself/itself?*] and the like), where *X* may be a person, as in (6a), or a thing, as in (7a).

Tuvaluan (Austronesian, Remote Oceanic, Nuclear Polynesian; Tuvalu)

- (5) a. *ko-oi tou fenua?*
 FOC-who 2SG.POSS island
 ‘What’s your home island?’ or ‘What’s your native country?’ (lit.: ‘Who (is) your island?’) (Besnier 2000:422)
- b. *ko-oi t-tino naa?*
 FOC-who DEF.SG-person that
 ‘[Addressed to someone in the dark:] Who is there? (lit.: ‘Who (is) that person?’)’ (Besnier 2000:424)

Semelai (Austro-Asiatic, Mon-Khmer, Aslian; Malaysia)

- (6) a. *kadeh glbr?*
 who be.named
 ‘What are (you) called?’, ‘What is your name?’ (Nicole Kruspe, p.c.)
- b. *kadeh na-təʔen?*
 who DEM-to.down
 ‘Who is the one (coming) down?’ (Kruspe 1999:293)

Hadendowa Bedawi/Beja (Afro-Asiatic, North Cushitic; Sudan)

- (7) a. *oo-tam ʔaab eedna t-ʔarabyeet-iib?*
 ART.M.SG.ACC-sorghum.ball who.ACC say.IPFV.3PL ART.F.SG-Arabic-in
 ‘What is *sorghum ball* in Arabic? (lit.: ‘Who do they call *sorghum ball* in Arabic?’)’ (Martine Vanhove & Mohamed-Tahir Hamid Ahmed,

p.c.)

b. *ʔaab irhitaa?*

who.ACC see.PFV.2SG.M

‘Who did you see?’ (Roper 1928:34)

The Hadendowa Bedawi example (8), literally meaning something like ‘Who do they call your name?’, represents a variation on the [IPW is X named?]-construction.

Hadendowa Bedawi/Beja (Martine Vanhove & Mohamed-Tahir Hamid Ahmed, p.c.)

(8) *oo-sm-ook ʔaab eedna?*

ART.M.SG.ACC-name-2SG.POSS.ACC who.ACC say.IPFV.3PL

‘What are you called? (lit.: ‘Who do you they call your name?’)’

Consider also the following Barasana construction in (9), where the word *~wábé* ‘name, thing, species, kind’¹² is incorporated into the verb *kítí-* ‘have’. Similarly to the Hadendowa Bedawi example (8), I prefer to consider (9) as a subtype of the [IPW is X named?]-construction.

Barasana (Eastern Tucanoan; Colombia; Elsa Gómez-Imbert, p.c.)

(9) *~jib-î ~wábé-kítí-á-ti bî?*

IPW.AN-AN.M.SG NON<3>.name-have-IPFV-Q 2SG

‘[To a man:] What are you called? (lit.: ‘Who do you have as a name?’)’

Languages may consistently show a strong preference for only one of the three constructions presented. They may also use different constructions and/or different resolution (or avoidance) strategies in different contexts. Often, the variation appears to depend on the semantics of the element *X* involved in the constructions [IPW is X’s name?], [IPW is X (thing)? (by name)] or [IPW is X named?]. In Hadendowa Bedawi, for instance, the construction [IPW is X named?] involving the verb *eedna* ‘say.IPFV.3PL’ and the interrogative *ʔaab* ‘who.ACC?’, as in (7a) or (10), is obligatory with any non-human *X*, whereas with a human *X* both [IPW is X named?] and [IPW is X’s name?]-constructions involving ‘who?’ are possible as well, as illustrated with (8) and (11) respectively.

¹² The tilde before Barasana forms marks nasalization, which is analyzed as a suprasegmental feature of the morpheme that can also spread to other morphemes.

Hadendowa Bedawi/Beja (Afro-Asiatic, North Cushitic; Sudan; Martine Vanhove & Mohamed-Tahir Hamid Ahmed, p.c.)

- (10) *oon w-hendi (oo-sm-oo)*
 DEM.M.SG.ACC ART.M.SG.ACC-tree (ART.M.SG.ACC-name-3SG.POSS.ACC)
ʔaab eedna?
 who.ACC say.IPFV.3PL
 ‘What do they call this tree? (lit.: ‘Who do you they call (the name of) this tree?’)’
- (11) *oo-sm-ook ʔaab-i?*
 ART.M.SG.ACC-name-2SG.POSS.ACC who.ACC-COP.3SG.M
 ‘What is your name? (lit.: ‘Who is your name?’)’

In Semelai, as illustrated in (6a) above, a person’s name is asked with the question ‘Who are you called?’. However, the name of a non-human entity, such as a tree in (12), is questioned with ‘How do you call it?’. Note also that different verbs are used in (12) and (6a).

Semelai (Nicole Kruspe, p.c.)

- (12) *dlɔŋ nɔʔ, hɔ̃nmande ji=cəl?*
 tree this how 2=say
 ‘What do you call this tree? (lit.: ‘This tree, how do you call?’)’

Languages may also allow for different resolution or avoidance strategies within the same construction involving one and the same kind of entity X, as in (13a) vs. (13b) from Nama.

Nama (“Khoisan”, Khoe-Kwadi, Khoekhoe; Namibia; Levi Namaseb & Wilfrid Haacke, p.c.)¹³

- (13) a. *sa /ʔon-s-a ʔa tare?*
 2 name-3SG.F-OBL COP what
 ‘What is your name?’
- b. *sa /ʔon-s-a ʔa tari?*
 2 name-3SG.F-OBL COP who
 ‘What is your name?’ (lit.: ‘Who is your name?’)

Similarly, a different resolution or avoidance strategy may be required (or allowed) with one and the same kind of entity X but within a different construction. Compare in this respect (13) with (14), where *mati* ‘how?’ must be used rather than *tare* ‘what?’ or *tari* ‘who?’.

¹³ Tones are not marked.

Nama (Levi Namaseb & Wilfrid Haacke, p.c.)

- (14) *sa-ts-a* *mati* *ra* *≠ai-he?*
 2-2SG.M-OBL how PRS.INCHOATIVE call-PASS
 ‘What are you called? (lit.: ‘How are you called?’)’

Note that a variation in resolution strategies within the same construction involving one and the same element X need not be free. In some languages, the use of different interrogatives may imply some (subtle) difference in meaning. Compare, for instance, Wulai Atayal examples (15a) and (15b).

Wulai Atayal (Austronesian, Atayalic; Taiwan; Huang 1996:293)

- (15) a. *ima?* *lalu?*=*su??*
 who name-2SG.GEN
 ‘What is your name?’ (lit.: ‘Who is your name?’)
 b. *nanu?* *lalu?*=*su??*
 what name-2SG.GEN
 ‘What is your name?’

According to Huang (1996:293), “[t]he difference between the two is that sentence [(15a)] is used when, for example, two people meet and one asks for the other’s name, while [(15b)] is used when, for instance, in the local government office, the actual writing of one’s name is being requested”.¹⁴ In other words, *ima?* ‘who?’ (15a) is used when the personal proper name asked about is viewed as a personal proper name proper, i.e. as being inseparably related to the person that bears it, so that the use of *ima?* ‘who?’ just highlights its categorical presuppositional meaning ‘person’. The interrogative *nanu?* ‘what?’ (15b) is used when the personal proper name asked about is viewed as a pure autonym whose categorical presuppositional meaning is not ‘person’ but ‘word’ or ‘name’, so that the question in (15b) is understood as a question about the form of the proper name.

Finally, what may look as one of the constructions discussed above does not necessarily involve the non-prototypical combination of values of the kind [thing + identification + proper name]. In this respect, compare (16a) and (16b) from Badiaranke/Badyara.

¹⁴ Interestingly, according to Huang (1996:293), no such distinction seems to exist in the closely related Mayrinax Atayal.

Badiaranke/Badyara (Niger-Congo, Northern Atlantic; Senegal; Rebecca Cover, p.c.)

- (16) a. *komitf-ẽ* *kõ* *wa: i?*
 name-2SG.POSS ART what COP
 ‘What is your name? (lit.: ‘Your name, what is (it)?’),¹⁵
- b. [A:] *komitf-ẽ* *kõ* *wa: i?*
 name-2SG.POSS ART who COP
 [B:] *ma:m-ã* (*mitf-e:n-õ*)
 grandparent-1SG.POSS call-APPL-PASS.3SG
 ‘[A:] Who are you named for? (lit.: ‘Your name, who is (it)?’) [B:] (I’m
 named for) my grandparent’

Question (16b) should probably be paraphrased as something like ‘Who is it (originally) with your name?’. Alternatively, if we analyze *komitf-ẽ* here not as ‘your name’ but as ‘your namesake’, (16b) can be paraphrased as ‘Who is your namesake?’. In any case, there is no non-prototypical combination of values here.

3.3 Contexts

In this section, I will present the kinds of contexts (i.e., the kinds of proper names) for which ‘who?’-dominance in the case of the non-prototypical combination of values [thing + identification + proper name] has been found: proper names of persons in (Section II.3.3.1), proper names of domestic animals (Section II.3.3.2), toponyms (Section II.3.3.3), temporal proper names (Section II.3.3.4), names of “folk genera” (species) (Section II.3.3.5), proper names of inanimate things (Section II.3.3.6), and pure autonyms (Section II.3.3.7).

3.3.1 Proper names of persons

Questions about proper names of persons clearly form the most common context where ‘who?’-dominance in the case of the non-prototypical combination of values [thing + identification + proper name] can be found in the languages of the world. The most common type of construction involved here appears to be the construction [Who is X’s name?], cf. examples (1), (11), (13b), (15a) above. Still, some languages prefer the construction [Who is X named?], as e.g. Semelai (6), Hadendowa Bedawi (8), and Barasana (9) above. The possibility to use ‘who?’ in questions about personal proper names also seems to be a prerequisite for the possibility of ‘who?’ in [Who is X’s name?] and [Who is X

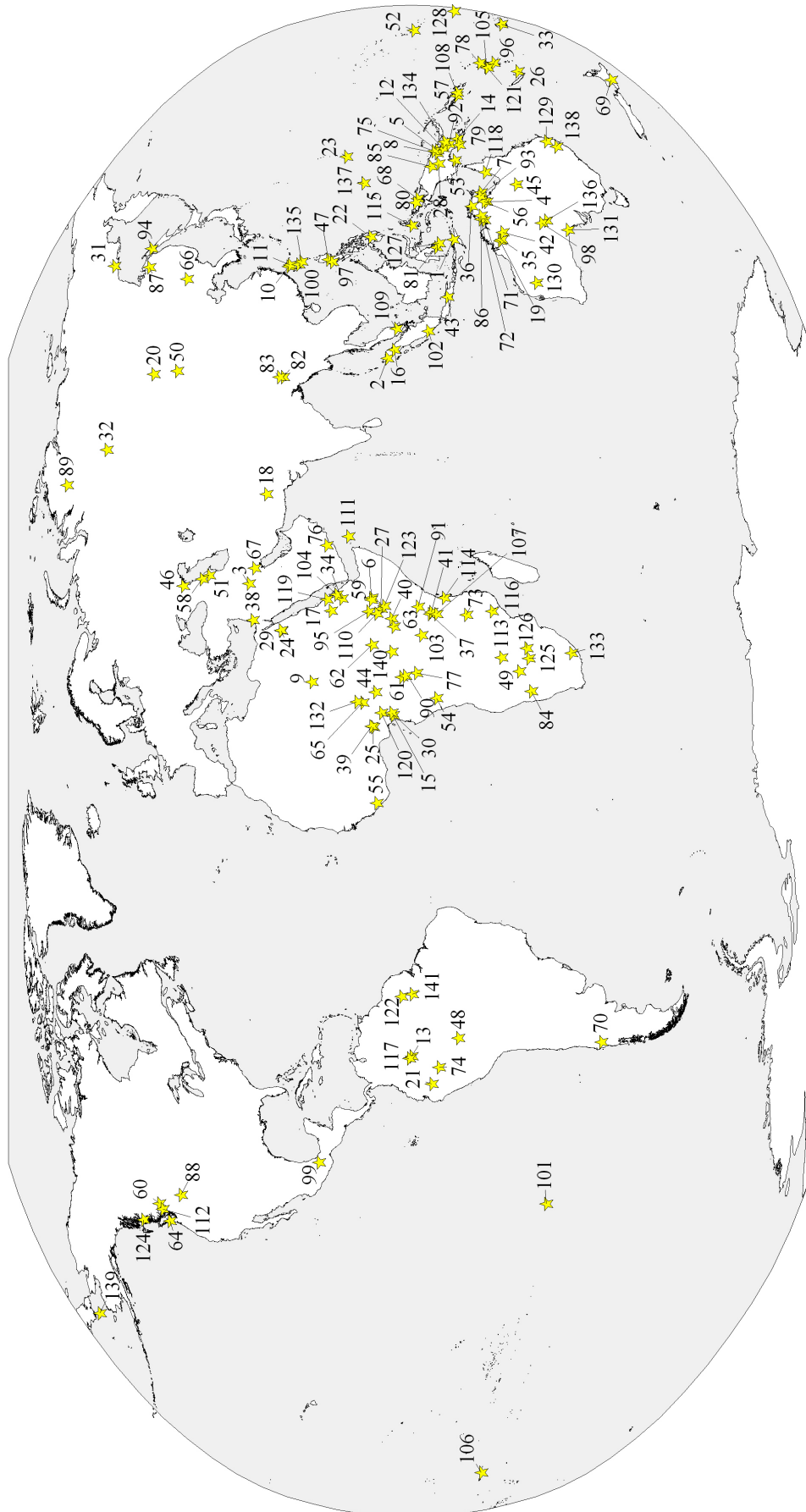
¹⁵ The question about someone’s name seems to be more commonly asked as *ana: mitf-o-i* (how call-PASS-2SG) ‘What are you called? (lit.: ‘How are you called?’) (Rebecca Cover, p.c.).

named?'] constructions with any other kinds of proper names. As discussed in Section II.3.1, most languages that prefer 'who?' in questions about personal proper names appear to do so because they take into account the categorical presuppositional meaning of personal proper names, viz. the fact that such names denote persons.

Map 1 illustrates the distribution of languages that have been found to allow for 'who?'-dominance in the case of the non-prototypical combination of values [thing + identification + proper name] in questions about proper names of persons. As can be observed, certain geographical areas and linguistic groups appear to be particularly prone to use 'who?' in questions about personal proper names. Thus, this phenomenon is particularly widespread in Oceania and island Southeast Asia in Austronesian languages, (non-Austronesian) Papuan languages (phyla such as Trans-New Guinea, West and East Papuan, Sepik-Ramu, East Bird's Head), and Australian languages (both Pama-Nyungan and non-Pama-Nyungan). In Africa and the Middle East, the relevant languages tend to concentrate in Eastern and Central Africa among Afro-Asiatic languages (especially in Cushitic, Omotic, Semitic and Egyptian branches), Niger-Congo languages (mainly Bantu), Nilo-Saharan (especially, Central and Eastern Sudanic). In Eurasia, most examples appear to be confined to Tungusic and Mongolian languages. In North America, we find the relevant examples mainly in the languages of the northwestern coastal region, viz. in families such as Salishan, Wakashan, Tsimshian and Plateau Penutian. In South America, the phenomenon in question appears to be found mostly in Carib, Tucanoan and Panoan languages. In other parts of the world and other linguistic groups, examples appear to be attested only sporadically.

At least in some cases, the high concentration of languages with 'who?'-dominance in the case of the non-prototypical combination of values [thing + identification + proper name] in questions about proper names of persons appears to correlate with an overall prominent status of proprial nominals in grammars of the respective languages. Thus, Austronesian languages typically have various "personal" articles or special "personal" case markers for personal proper names (cf., e.g., Section III.4.2.2.1.2). In descriptions of many Oceanic Austronesian languages, proper names are typically set apart as a distinct subclass of nominals definable on purely morphosyntactic grounds. In Pama-Nyungan languages of Australia, proper names often differ from common nominals in their case marking (cf., e.g., Section III.6.4.1.1, especially Sections III.6.4.1.1.1 and III.6.4.1.1.4). Somewhat similarly, in many older languages of the Semitic branch of Afro-Asiatic, proper names seem to have differed from common nouns as regards the marking of case and the so-called "state" (cf., e.g., Lipiński 1997:258-259, 262-265, 273).

Map 1. The distribution of languages allowing ‘who?’ in NAME-questions about personal proper names



Language	Family (Ethnologue)	Language	Family (Ethnologue)	Language	Family (Ethnologue)
1 Abui	Trans-New Guinea	24 Coptic	Afro-Asiatic	47 Karao	Austronesian
2 Aceh	Austronesian	25 Dan (Gweetawu)	Niger-Congo	48 Kaxarari/Kaxariri	Panoan
3 Akkadian	Afro-Asiatic	26 Dehu/Drehu	Austronesian	49 Kgalagadi	Niger-Congo
4 Alawa	Australian	27 Dime	Afro-Asiatic	50 Khalkha/Halh	Altaic
5 Amele	Trans-New Guinea	28 Duna	Trans-New Guinea		Mongolian
6 Amharic	Afro-Asiatic	29 Egyptian (pre-Coptic)	Afro-Asiatic	51 Khinalug	North Caucasian
7 Anindilyakwa	Australian	30 Eton	Niger-Congo	52 Kiribati	Austronesian
8 Apali	Trans-New Guinea	31 Even	Altaic	53 Kiwai (Southern, Iasa)	Trans-New Guinea
9 Arabic (Chadian)	Afro-Asiatic	32 Evenki (Western, Poligus)	Altaic	54 Kongo (San Salvador)	Niger-Congo
10 Atayal (Mayrinax)	Austronesian			55 Kpelle	Niger-Congo
11 Atayal (Wulai)	Austronesian	33 Fijian	Austronesian	56 Kriol (Roper River area)	Creole
12 Awara	Trans New-Guinea	34 Geez	Afro-Asiatic	57 Lavukaleve	East Papuan
13 Barasana	Tucanoan	35 Gooniyandi	Australian	58 Lezgian	North Caucasian
14 Baruga (Tafota)	Trans-New Guinea	36 Gunwinggu/Mayali	Australian	59 Libido	Afro-Asiatic
15 Basaa (Mbene)	Niger-Congo	37 Hadza/Bali	Khoisan	60 Lillooet	Salishan
16 Batak (Karo)	Austronesian	38 Hebrew (Ancient)	Afro-Asiatic	61 Lingala	Niger-Congo
17 Beja/Bedawi	Afro-Asiatic	39 Idoma	Niger-Congo	62 Luwo	Nilo-Saharan
		40 Ik	Nilo-Saharan	63 Ma'di	Nilo-Saharan
18 Brahui	Dravidian	41 Iraqw	Afro-Asiatic	64 Makah	Wakashan
19 Bunaba/Bunuba	Australian	42 Jaru/Djaru	Australian	65 Mambay	Niger-Congo
20 Buriat (Russia)	Altaic	43 Javanese	Austronesian	66 Manchu	Altaic
21 Candoshi	Isolate	44 Kab(b)a (of Paoua)	Nilo-Saharan	67 Mandaic/Neo-Mandaic	Afro-Asiatic
22 Cebuano	Austronesian	45 Kalkutung/Kalkatungu	Australian	68 Manikion/Sough	East Bird's Head
23 Chamorro	Austronesian	46 Kalmyk	Altaic	69 Maori	Austronesian

Language	Family (Ethnologue)	Language	Family (Ethnologue)	Language	Family (Ethnologue)
70 Mapuche/Mapudungun	Araucanian	93 Nunggubuyu	Australian	114 Swahili	Niger-Congo
71 Marringarr/Maringarr	Australian	94 Orok	Altaic	115 Taba/Makian (East)	Austronesian
72 Marrithiyel/Marithiel	Australian	95 Oromo (West Central, Mecha)	Afro-Asiatic	116 Takwane	Niger-Congo
73 Matengo	Niger-Congo	96 Paama/Paamese	Austronesian	117 Tatuvo	Tucanoan
74 Matsés	Panoan	97 Pangasinan	Austronesian	118 Thayore/Kuuk	Australian
75 Mauwake	Trans-New Guinea	98 Pitjantjatjara	Australian	Thaayorre	
76 Mehri/Mahri	Afro-Asiatic	99 Popoluca (Sierra)	Mixe-Zoque	119 Tigre	Afro-Asiatic
77 Mongo-Nkundo	Niger-Congo	100 Puyuma	Austronesian	120 Tikar	Niger-Congo
78 Motlav/Mwotlap	Austronesian	101 Rapa Nui	Austronesian	121 Tolomako	Austronesian
79 Motu	Austronesian	102 Rejang	Austronesian	122 Trió	Carib
80 Mpur	West Papuan	103 Rwanda	Niger-Congo	123 Tsamai/Ts'amakko	Afro-Asiatic
81 Muna	Austronesian	104 Saho (Asa-Awurta & Tarua)	Afro-Asiatic	124 Tsimshian (Coast)/	Penutian
82 Naga (Angami)	Sino-Tibetan	105 Sakao	Austronesian	Sm'algyax	
83 Naga (Ao)/Ao	Sino-Tibetan	106 Samoan	Austronesian	125 Tswana	Niger-Congo
84 Nama (Khoekhoe)	Khoisan	107 Sandawe	Khoisan	126 Tswapong	Niger-Congo
85 Namia	Sepik-Ramu	108 Savosavo	East Papuan	127 Tukang Besi	Austronesian
86 Nangikurunggurr/Ngankikurungkurr	Australian	109 Semelai	Austro-Asiatic	128 Tuvaluan	Austronesian
87 Negidal (Nizovskiy)	Altaic	110 Sheko	Afro-Asiatic	129 Wakawaka	Australian
88 Nez Perce	Penutian	111 Soqotri	Afro-Asiatic	(Duungidjauu)	
89 Nganasan (Ust'-Avam)	Uralic	112 Squamish	Salishan	130 Watjarri/Wajarri	Australian
90 Ngombe (Ligenza)	Niger-Congo	113 Subiya/Kuhane/Ciikuhane	Niger-Congo	131 Wirangu	Australian
91 Nubi	Creole			132 Wuzlam	Afro-Asiatic
92 Numbami	Austronesian			133 Xhosa	Niger-Congo
				134 Yabem	Austronesian

Language	Family (Ethnologue)
135 Yami	Austronesian
136 Yankunytjatjara	Australian
137 Yapese	Austronesian
138 Yugambal-Bandjatang	Australian
139 Yup'ik (Central Siberian, Eskimo-Aleut St. Lawrence Island)	
140 Zande	Niger-Congo
141 Zo'é	Tupi

In Bantu languages, proper names (and other nominals that are not in need of referential disambiguation) often differ from the rest of the nominals in the agreement patterns they trigger and peculiarities in the use of the so-called “augment”¹⁶ with them, so that in some Bantu languages they may even be characterized as “genderless”, i.e. as being outside of the gender system (cf. Van de Velde 2006).

Furthermore, in some cases we may speak of an areal influence. Thus, Bantu languages seem to have influenced “Khoisan” languages. A somewhat clearer example of areal influence seems to be represented by Nganasan, the only Uralic language I know of to require ‘who?’ in questions about personal proper names, as in (17), which in all probability is due to Tungusic (viz. Evenki) influence.

Avam Nganasan (Uralic, Samoyedic; Russia; Valentin Goussev, p.c.)

- (17) *sjili/*maa n'im-ti?*
 who/*what name-3SG
 ‘What is his name? (lit.: ‘Who is his name?’)’

The Cushitic (and probably also Omotic) branch of Afro-Asiatic appears to represent an interesting example of the development of ‘who?’-dominance in questions about personal (and other) proper names not related to a prominent grammatical status of proper names. Rather, it seems to be due to a concurrence of certain developments in the evolution of the interrogative pronominal systems of these languages. Thus, as discussed in Section III.2.3.5.1 (cf. also Section III.2.3.4.1.1), ‘who?’ interrogatives in Cushitic languages seem to have developed out of an earlier selective interrogative pronominal ‘which one?’, which in turn, normally seems to go back to a locative interrogative proadverbial ‘where?’. At the same time, it is also not uncommon for modern Afro-Asiatic languages to use these two interrogatives as avoidance strategies in questions about (personal) proper names, as was illustrated with (3) from Hausa (West Chadic) and (4) from Amharic (West Semitic). Hence, it seems reasonable to suppose that the recurrence to ‘who?’-dominance in the case of the non-prototypical combination of values [thing + identification + proper name] in questions about personal (and other) proper names in Cushitic is due to the fact that (i) modern ‘who?’ interrogatives used to be ‘where?’ and/or ‘which one?’ interrogatives and that (ii) the latter interrogatives were also resorted to as avoidance strategy in questions about proper names. When the shift from ‘where?’ and/or ‘which one?’ to ‘who?’ occurred, Cushitic languages must have kept on using these interrogatives in questions about proper names. This hypothesis also offers an explanation for the fact that in some Cushitic languages ‘who?’-dominance appears to be found with any kind of proper names, at least

¹⁶ In terms of Greenberg (1978), augment is a “stage II” article.

within the [Who is X named?'] construction. Recall, for instance, the Hadendowa Bedawi examples (7a), (8), (10), (11) provided above.

In fact, Cushitic languages may not be unique. Thus, the large range of contexts where several Bantu languages allow 'who?'-dominance in NAME-questions may also be due to a similar evolution of an originally locative-selective interrogative to 'who?'. In this respect, see the discussion in Section III.2.1.1.2.3. The overall prominent status of proprial nominals in grammars of Bantu languages mentioned above may have further helped the development of 'who?'-dominance in NAME-questions.

3.3.2 *Proper names of domestic animals*

In many languages, it appears possible, in an appropriate context, to ask about the proper names of a domestic animal simply by using the interrogative 'who?' in the same kinds of syntactic constructions where 'who?' would be used to ask about a human, as in (18) from Dutch.

Dutch

- (18) [A:] *Wie heeft er zo veel melk gegeven?* [B:] *Bella*
 who has there so much milk given PROP
 '[A person who lives on a farm and knows the cows personally:] Who has
 given so much milk? [B:] Bella (the name of a cow)']

Questions like (18) seem to be best treated as instances of personification, rather than non-prototypical combinations of values proper. This interpretation seems to be supported by the fact that the use of 'who?' in questions about proper names of domestic animals seems to be rather widespread, also in languages such as Dutch, where otherwise, animate things (including animals) are questioned with 'what?' rather than 'who?'. I have not tested this consistently but my impression is that the only prerequisite for such questions to be possible with 'who?' is that a given linguistic community has domestic animals and that domestic animals are important enough to be given proper names.

However, there are also languages that allow (or require) the use of 'who?' in questions about proper names of domestic animals in constructions [IPW is X's name?] and/or [IPW is X named?'] just as they allow (or require) the use of 'who?' in the same constructions in questions about proper names of persons, as in (19) from Angami Naga.

Angami Naga (Sino-Tibetan, Tibeto-Burman, Kuki-Chin-Naga; India;
 Giridhar 1980:36)

- (19) a. *ñ zâ šūpuò gā?*
 your name who Q
 'What is your name?' (lit.: 'Who is your name?')

b. *n̂ fə zâ šūpuò gā?*

your dog name who Q

‘What is the name of your dog?’ (lit.: ‘Who is the name of your dog?’)

Although the reason for the use of ‘who?’ in (19b) is also clearly the assimilation of domestic animals to humans in some respect (i.e., their personification),¹⁷ we are dealing here with a non-prototypical combination of values of the kind [thing + identification + proper name], because it is the identity of a proper name that is being questioned. Therefore, the possibility of examples such as (19b) should be mentioned explicitly.

Since I have not looked for the possibility of questions like (19b) consistently, few generalizations can be made here. However, I have a strong impression that necessary and sufficient conditions for [Who is X’s name?] or [Who is X named?] about domestic animals to be possible in a given language are as follows: (i) it must be possible to use [Who is X’s name?] or [Who is X named?] for humans and (ii) it must be possible to ask questions like (18) (i.e., it must be possible to personify domestic animals).

3.3.3 Proper names of places (toponyms)

Some languages have been found to allow for questions about proper names of places to be constructed as [Who is X?], as can be illustrated with (20a) from Tuvaluan.

Tuvaluan (Austronesian, Remote Oceanic, Nuclear Polynesian; Tuvalu)

(20) a. *ko-oi tou fenua?*

FOC-who 2SG.POSS island

‘What’s your home island?’ or ‘What’s your native country?’ (lit.: ‘Who (is) your island?’) (Besnier 2000:422)

b. *ko-oi t-tino naa?*

FOC-who DEF.SG-person that

‘[Addressed to someone in the dark:] Who is there? (lit.: ‘Who (is) that person?’)’ (Besnier 2000:424)

According to Besnier (2000:422), in Tuvaluan, interrogative *ai/ei/oi* ‘who?’ is “also used in questions about names of places of regional or political import. Islets, islands and atolls, island groups, countries, and continents fall in this category”. Together with personal proper names, these words form a lexico-

¹⁷ According to Giridhar (1980:36), *šūpuò* ‘who?’ will be used here because “names of human beings and dogs are considered human”. Note in this respect that “names of human beings and animals like ‘dog’ [also] take case markers for the human noun” (Giridhar 1980:26).

syntactic class of proprial nominals in Tuvaluan. As can be readily observed, not all place-names belong to the latter class in Tuvaluan and respectively not all place-names can be questioned with ‘who?’. My impression is that what matters here is not “regional or political import” but rather the conceptualization of the respective places as inhabited (or at least, as typically inhabited), that is as related to humans.¹⁸ Note in this respect that Besnier (2000:371) glosses the word *fenua*, which in (20a) was translated with ‘island’ or ‘country’, as ‘*island community*, island, country’ (emphasis added), cf. (21).

Tuvaluan (Besnier 2000:371)

- (21) [*Koo kkese te faiga o te feagaiga*,]
koo see gaalue te fenua, [*a koo ttofi fua se potu kau*]
 INCEPTIVE NEG work DEF.SG island.community
 ‘[The contract’s terms have changed,] the island community no longer
 does the work, [but (instead) a (work) team is appointed]’

In examples such as (22) and (23), the original implication ‘inhabited’ becomes even more prominent so that in fact, the proper names here do not refer to the inhabited places as such anymore but rather, metonymically, to the people that inhabit these places. Note in this respect that the English translations of (22) and (23) also (allow to) use ‘who?’. Obviously, this is neither a non-prototypical combination of values nor a lack of differentiation between ‘who?’ and ‘what?’.

Tuvaluan

- (22) *Falani ne takavale ia ai i t-taua teenaa?*
 France PST defeated at who at DEF.SG-war that
 ‘Who defeated France in that war?’ (Besnier 2000:550)
- (23) *A Niu Kaletonia e puke nee ei?*
 ABS.CONTRASTIVE New Caledonia NON<PST> hold ERG who
 ‘Which country does New Caledonia belong to?’ (Besnier 2000:423) or

¹⁸ In this respect, compare the discussion in Van Langendonck (2007:207-218). Thus, Van Langendonck (2007:207) notes that “among the parameters that could apply [for distinguishing “semantic place name categories”], we especially count the nature of the soil and the *human interaction* [emphasis added]. [...] As for the human interaction with these places, there are such features as inhabitable or not, administratively structured and bounded, habitat, dwelling-place vs. vaster inhabitable area, familiar, integrated, i.e. human-friendly vs. isolated, remote or elevated”. For instance, “[i]n Germanic languages and also in Spanish, the occurrence of the definite article in place names is apparently determined by the absence of human organization or administration”, so that, for instance, “[w]hen regions develop into states, the article seems to disappear, at least in English and Dutch”, as in *the Lebanon* → *Lebanon* or *the Ukraine* → *Ukraine* (Van Langendonck 2007:208).

‘Who owns New Caledonia?’

Still, it seems reasonable to consider at least examples such as (20a) as instances of the non-prototypical combination of values of the kind [thing + identification + proper name]. In all probability, there must be more Oceanic (or perhaps even Austronesian) languages of this kind.

Another language where it appears possible to construct questions about proper names of (at least some) places as [Who is X?] is Ngombe (24a).

Ngombe (Niger-Congo, Bantu C50; Democratic Republic of the Congo; Rood 1958:xxi)¹⁹

(24) a. *ngando íyě nda?*

G9.village AG9.this who

‘What is the name of this village? (lit.: ‘Who is this village?’)

b. *moto íyǒ nda?*

G1.man AG1.this who

‘Who is this man?’

Compare also Bantu languages that allow (or require) for questions about proper names of places to be constructed as [Who is X’s name?] and/or [Who is X named?] in (27) below. I do not know whether the same is possible in Ngombe.

Besides Tuvaluan and Ngombe, at least three Australian Pama-Nyungan languages, viz. Jaru/Djaru (WJa2;²⁰ Tsunoda 1995), Pintupi (WDf; Hansen & Hansen 1978), and Wajarri/Watjarri (WGa1; Douglas 1981), and one East Papuan language of Solomon Islands, viz. Lavukaleve (Terrill 2003), may prove to allow for questions about toponyms to be constructed as [Who is X?]. (In the three Pama-Nyungan languages, the respective ‘who?’ interrogative has the same form *ngana-* /ɲana/). Unfortunately, for Wajarri, no examples are provided and all the relevant examples provided in the sources on the remaining languages are not interrogative in the strict sense. Rather, in Pintupi, as discussed in Section III.6.4.1.1.4, the “interrogative” in example (III.6.4.1.1.4:29) seems to be used as a filler, i.e. a word like ‘whatchamacallit’ or ‘whatsit’. The same appears to be the case in the Lavukaleve example (25).

¹⁹ The variety of Ngombe described by Rood (1958) is Ngenja, or Ligenza in Ethnologue’s denomination.

²⁰ Cf. Map III.6.1:2.

Lavukaleve (East Papuan; Solomon Islands)

- (25) *lai vo-e-me-ge lai vo-e-me-ge*
 paddle 3PL.OBJ-SUB-HAB-ANTERIOR paddle 3PL.OBJ-SUB-HAB-ANTERIOR
aka / mi / ami hi? / Sokopiu
 then um who 3SG.N.FOC PROP
 ‘They paddled and paddled and paddled, then – where? – Sokopiu (a place name)’ (Terrill 2003:494), or maybe better: ‘... (and) then (they arrived in) ... um... whatitsname... Sokopiu’)

In Jaru, in all examples where the “interrogative” *ɲana-* is used about a place name, it is marked with the clitic *-wari* ‘it is not known...’, as in (26).²¹ Note that (26b) also involves a verb ‘call, name’.

Jaru (Pama-Nyungan; Australia)

- (26) a. *ɲana-wari*
 who-I.WONDER
 ‘[(From a narrative:) We said thus: ‘Oh, this water lies in (a place) with no creek’.] I wonder/ I don’t know what is the name (of this)’ (Tsunoda 1995:253)
- b. *ɲana-wari-lu maran-an binga jambi*
 who-I.WONDER-3PL.NOM call-PRS creek big
 ‘[(From the same narrative as the preceding example) We went east, across sight to a big river.] I don’t know what they call the big river’ (Tsunoda 1995:257)²²

At least in Wajarri and Pintupi, place-names appear to form a separate morphosyntactic subclass of nominals together with personal proper names.²³

Furthermore, I have found several languages that allow (or require) for questions about proper names of places to be constructed as [Who is X’s name?] and/or [Who is X named?]. These languages are summarized in (27). I do not know whether constructions [Who is X’s name?] and/or [Who is X named?] about proper names of places are also possible in the languages discussed earlier

²¹ In Jaru, *ɲana* also seems to be used in questions about the names of folk genera (cf. Section II.3.3.5).

²² There is another very similar example (Tsunoda 1995:254).

²³ Tsunoda (1995) does not mention any morphosyntactic peculiarities of Jaru proper names, either personal or locative. In Lavukaleve, “proper names referring to places” are subdivided on morphosyntactic grounds into “place nouns”, such as *Sokopiu*, and “locational nouns”, such as *Karumulu* (the name of a village in the central Russells), both being different from “common nouns” (Terrill 2003:44, 165-166). Note that the Lavukaleve class of “common nouns” includes “many of what are traditionally termed proper names” (Terrill 2003:44).

in the present section. I have a more or less appropriate example only for Jaru, cf. (26b) above.

(27) Africa

Niger-Congo, Bantu:

Tswana (S30; Botswana & South Africa; Andy Chebanne, p.c.)

Tswapong (S30; Botswana; Bennett Maifala, p.c.)

Kgalagadi (S30; Botswana; Kems Monaka, p.c.)

Subiya/Kuhane (K50; Namibia & Botswana; Ndana Ndana, p.c.)

Takwane (P30; Mozambique; Oliver Kröger, p.c.)

Afro-Asiatic, Cushitic:

Hadendowa Bedawi (North Cushitic; Sudan; Martine Vanhove & Mohamed-Tahir Hamid Ahmed, p.c.)

Libido (Highland East Cushitic; Ethiopia; Joachim Crass, p.c.)²⁴

West Central Oromo (East Cushitic; Ethiopia; Tolemariam Fufa, p.c.)

Afro-Asiatic, West Semitic:

Amharic (Ethiopia; Joachim Crass, p.c.; Seyoum Mulugeta, p.c.)²⁵

Nilo-Saharan, Central Sudanic, East:

Ma'di (Uganda; Mairi J. Blackings, p.c.)

Eurasia

Altaic, Northern Tungusic:

Poligus Evenki (Russia; Konstantinova 1968:73)

Southeast Asia & Oceania

Austronesian, Remote Oceanic, Nuclear Polynesian:

Rapa Nui (Viki Haoa Cardinali, p.c.)

New Guinea

Trans-New Guinea, Main Section, Central & South New Guinea:

Duna (Papua New Guinea; Lila San Roque, p.c.)²⁶

²⁴ There is some inter-speaker variation as regards the use of 'who?' or 'what?' here.

²⁵ There is some inter-speaker variation as regards the use of 'who?' or 'what?' here.

²⁶ Here, some speakers prefer 'who?', but most seem to favour 'what?'.

Australia

Pama-Nyungan:

?Jaru (Ngumpin-Yapa; WJa2;²⁷ Tsunoda 1995; cf. above)?Pitjantjatjara (Western Desert/Wati; WDM; Eckert & Hudson 1988)²⁸

?Yankunytjatjara (Western Desert/Wati; WDn; Goddard 1983 via Mushin 1995; Goddard 1994)

non-Pama-Nyungan, Bunaban.²⁹

?Bunuba/Bunaba (NF1; McGregor 2004:129)

?Gooniyandi (NF2; McGregor 2004:129)

By way of illustration, consider (28) from Kgalagadi and (29) from West Central Oromo.

Kgalagadi (Niger-Congo, Bantu S30; Botswana; Kems Monaka, p.c.)

(28) [A:] *libizho la lehele lo ke anye?*

G5.name AG5.CON G11.place AG11.DEM COP who

[B:] *ke Huhunsi*

COP PROP

‘[A:] What is the name of this place? [B:] It’s Hukuntsi (a village name)’

West Central Oromo (Afro-Asiatic, East Cushitic; Ethiopia)

(29) [A:] *mak’aan bakka kanaa eeñu?*

name.NOM place this.GEN who

[B:] *Addis Ababa-dā*

PROP-COP

‘[A:] What is the name of this place? [B:] It’s Addis Ababa (the capital of Ethiopia)’ (Tolemariam Fufa, p.c.)

In most cases, the languages in (27) do not seem to impose any particular

²⁷ Cf. Map III.6.1:2.

²⁸ The situation in Pitjantjatjara and Yankunytjatjara is not completely clear, since in the relevant examples available to me, the interrogative *ngana-*, elsewhere used as ‘who?’, seems to be used only attributively to *ngura* ‘place’ or *ini* ‘name’. Alternatively, *ngana* might be in appositional relation to these nominals here, but this seems to be contradicted by statements such as the following: “[in Yankunytjatjara *ngana-* can] also be used to ask about the name of a place, e.g. *ngura nganala?* (literally, ‘place ‘who’-LOC’) ‘at what (named) place?’ [, however] used alone, i.e. without *ngura* ‘place’, it always means ‘who’, never ‘what named place’” (Goddard 1994:259).

²⁹ McGregor (2004:129) does not provide the forms of the respective interrogatives, let alone examples of their use. I have not found any relevant examples in McGregor’s (1990) description of Gooniyandi either and I have not been able to check Rumsey’s (2000) description of Bunuba.

restrictions on the kinds of places involved, the only condition is that the place is named (i.e., it is not just the name of a form of landscape, such as *swamp* or *mountain*). However, at least for one language, viz. Poligus Evenki, names of rivers are reported by Konstantinova (1968:73) to be the only kind of toponyms that can be questioned with ‘who?’, as in (30).

Poligus Evenki (Altaic, Northern Tungusic; Konstantinova 1968:73)

- (30) *ər bi:ra ɲi: gərbi:-n?*
 this river[NOM] who name-3SG.POSS
 ‘What (lit.: ‘who?’) is the name of this river?’

As far as the reasons for the use of ‘who?’ in questions about proper names of places are concerned, at least the following can be said. It seems that for Bantu and Pama-Nyungan languages, this use can be explained by appealing to the same prominent morphosyntactic status of proper names as was already discussed in Section II.3.3.1.³⁰ Furthermore, for Bantu an additional explanation appealing to the peculiarities of the evolution of their interrogative pronominals is possible (cf. Section II.3.3.1). Also see Section II.3.3.1 for a similar explanation of the situation in Cushitic languages. In Amharic and Ma’di, the use of ‘who?’ in questions about toponyms may be due to Cushitic influence. Similarly, in Australia, the presumed non-Pama-Nyungan examples from the Bunaban family may be due to Pama-Nyungan influence.

3.3.4 Temporal proper names

According to Van Langendonck (2007:225-231), temporal names indicating points or periods in time, such as *Monday* or *May*, can function as proper names. What is more, for some kinds of temporal names the proprial function is more typical than for others, which implies that it is more typical for some kinds of temporal names than for others to be construed as proprial lemmas. Thus, Van Langendonck (2007:231) argues that in English (and Dutch) “[t]he names of months make use of proprial lemmas”, while so-called “proprio-appellative lemmas underlie the names of seasons and days”, for instance.³¹

Similarly to proper names of places discussed in Section II.3.3.3, some languages have been found to allow for questions about temporal proper names to be constructed as [Who is X?]. Clear examples are found only for Austronesian languages of the Nuclear Polynesian subgroup, viz. Tuvaluan (31)

³⁰ Thus, at least in the Bantu S30 languages in (27) proper names of places appear to trigger the same agreement patten AG1 as personal proper names. In Western Desert languages, proper names of places take the same case marking as personal proper names.

³¹ On the terms “proprial lemma”, “proprio-appellative lemma”, cf. footnote I.2.5:20.

and Rapa Nui (32).³²

Tuvaluan (Austronesian, Nuclear Polynesian, Samoic-Outlier; Tuvalu; Besnier 2000:423)

- (31) a. *ko-oi te maasina e fano ei koe?*
 FOC-who DEF.SG month NON<PST> go ANAPHORIC 2SG
 ‘What month are you leaving in?’ (lit.: ‘Who is the month you are leaving?’)
- b. *ko te maasina o oi e fano ei koe?*
 FOC DEF.SG month of who NON<PST> go ANAPHORIC 2SG
 ‘What month are you leaving in?’ (lit.: ‘It is the month of who that you are leaving?’)

Rapa Nui (Austronesian, Nuclear Polynesian, East; Chile; Du Feu 1996:22)

- (32) *ko-ai te ava’e ko tara hao hai vanaga tire?*
 FOC-who ART month FOC January INS language Chile
 ‘What (lit.: ‘who?’) is January in Spanish?’³³

The names of months appear to be the only kind of temporal names that can be questioned with ‘who?’ in Tuvaluan and Rapa Nui. Apparently, this is due to the fact that the names of months are the only kind of temporal names that belongs to the special morphosyntactic class of proper names. In Rapa Nui, for instance, this can be observed from the presence of the “personal” marker *a* with the names of months, such as *i a hora iti* ‘in August’ (lit.: ‘in PERSONAL August’), but not for instance, with the names of days, such as *i te mahana piti* ‘on Tuesday’ (lit.: ‘in SPECIFIC Tuesday’), or holidays, such as *i te Noere* ‘at Christmas’ (lit.: ‘in SPECIFIC Christmas’) (cf. Du Feu 1996:129-134). In other words, of all the temporal names only the names of months appear to make use of proprial lemmas.³⁴ Note the remarkable parallel in this respect with the behaviour of

³² However, in all probability, there must be more Oceanic (or perhaps even Austronesian) languages of this type.

³³ Admittedly, strictly speaking, this example is not a question about the identity of a proper name indicating a point or period in time, since the proper name *tara hao* ‘January’ is used here as a pure autonym, viz. ‘the word for *January*’ (cf. Section II.3.3.7). I cite it here for want of a better example. This is justified because in Rapa Nui, not all autonyms, it seems, can be questioned this way, but only those that are autonyms of nominals belonging to the morphosyntactic class of proper names, viz. proper names of persons, some places and months.

³⁴ Besides the names of months, only ‘summer’, as in *i a hora* ‘in summer’ (lit.: ‘in PERSONAL summer’), is a proper name as well. It should also be mentioned that only the names of months of Rapa Nui origine are proprial. Rapa Nui has two additional systems of month names, one using Spanish month names, such as *i febrero* ‘in February’, and one using English month

English (and Dutch) temporal names, as mentioned earlier in the present section.

I have not looked specifically for questions about temporal proper names in the languages of the world. However, I suppose that in languages such as Tswana and especially Hadendowa Bedawi, where questions about all or almost all kinds of proper names can be constructed as [Who is X's name?] and/or [Who is X named?], the same should also be possible for temporal proper names as well.

3.3.5 Names of “folk genera” (*species*)

As discussed by Wierzbicka (1996:351-375), on the basis of linguistic evidence, we can say that the concepts denoting living things (unlike those denoting artefacts) are organized following a “transitive hierarchy of categories, which is unique to the realm of living things (although in that realm, too, there are many concepts which are not included in that transitive hierarchy, such as, for example, *bush*)” (1996:374). Thus, the highest level of this hierarchy is composed of “life forms”, such as *tree* or *animal*, followed by “folk genera”, such as *oak* or *dog*,³⁵ sometimes further subdivided into “subgenerics”, such as *spaniel* or *poodle* in the case of *dog*. The lowest level in the hierarchy is occupied by “specific taxa, normally represented by secondary lexemes, such as *Siamese cat* or *blue spruce*” (1996:365). Not all living things belong to such a transitive hierarchy, though. In particular, Wierzbicka thinks here of “unaffiliated generics” (or “isolates”), such as *spider* or *ant*, (1996:361) and “quasi-life forms”, such as *bush* or *grass* (1996:365, 375).

Wierzbicka (1996:370) further argues that “folk genera [should be] seen as a kind of ‘proper name’ (in fact the ‘real name’ [...]) which cannot be reduced to a set of observable properties and which is linked with a presumed ‘underlying essence’”. Normally, however, names of folk genera,³⁶ such as *oak* and *sparrow*, are not described as proper names (cf., e.g., Van Langendonck 2007:34-36). Indeed, on strictly morphosyntactic grounds, they do not normally seem to

names, such as *i feruari* ‘in February’ (Du Feu 1996:131). These month names are not proprial.

³⁵ According to Wierzbicka (1996:358, 361) the most “operationally useful [...] criterion” for distinguishing life forms from folk genera “is that of [“linguistically recognized”] ‘polytypicity’: a life form is a category [of living things] which is thought of as comprising many different (named) kinds of entities”. Another important “possible difference between folk generic concepts and life form concepts is that the former – but perhaps not the latter – imply a ‘hidden nature’ or an ‘underlying essence’ which cannot be reduced to any observable attributes” (Wierzbicka 1996:367).

³⁶ The term *folk genera* is better here than *species* because it accentuates the fact that we are dealing not with a scientific classification in terms of *species* but with a “naive” ethnobiological classification as manifested in our everyday language use.

function as proper names, and therefore, are not construed as proprial or at least proprio-appellative lemmas. Obviously, Wierzbicka's characterization of folk genera "as a kind of 'proper name'" is based on purely semantic grounds. That is, it appears that for her, folk genera are a kind of proper name because according to her, they are semantically comparable to proper names in some respect, viz. in that they "imply a 'hidden nature' or an 'underlying essence' which cannot be reduced to any observable attributes" (Wierzbicka 1996:367). Still, unlike proper names proper, "folk genera can be fully defined, and [...] in this respect they do not differ from the names of cultural kinds (or any other words)" (Wierzbicka 1996:368).

Given that names of folk genera can be conceived as semantically similar to proper names, it is not surprising that in many languages, it is possible to ask about names of folk genera in the same way as one would ask about names of persons or places, cf. (33).

English

- (33) a. [A:] *What's the name of that guy?* [B:] *It's John*
 b. [A:] *What's the name of the city we visited last summer?* [B:] *It was San Francisco*
 c. [A:] *What's the name of that big tree in front of your house?* [B:] *It's a sycamore*

Note that the situation is often somewhat different with subgenerics (34a), specific taxa (34b) or life forms (34c).

English

- (34) a. [A:] *What's the name of this dog?* [B:] *It's Fido/ [?]a poodle*
 b. [?][A:] *What's the name of this spruce?* [B:] *It's a blue spruce*
 c. [?][A:] *What's the name of this?* [B:] *It's a bird*

Furthermore, if there are languages that use 'who?' in questions about various kinds of proper names, we may also expect that among these languages there will be some languages that treat names of folk genera similar to proper names by using 'who?' in questions about them as well. Indeed, several languages of this kind have been found.

Before proceeding further, it should be pointed out that I will not consider here languages such as Russian (Indo-European, Slavic) or Swahili (Niger-Congo, Bantu), where living things, at least those that are also grammatically animate, are normally questioned with 'who?' anyway (cf. Section II.4), even when it is not so much the identity of the name of a folk genus as such that is at issue but the classification of a living thing, as in 'What bit you?'. The reason is that if in such a language a personal proper name is questioned with 'who?', the

same appears to be necessarily the case for (grammatically animate) names of folk genera as well. In this respect, consider examples in (35-37) from Swahili. Note that the difference between (36) and (37) is due to the fact that nominals denoting trees are not grammatically animate in Swahili.

Swahili (Niger-Congo, Bantu G40; Tanzania, Kenya, etc.; Assibi Amidu, p.c.)

- (35) a. *huyu ni nani?*
AG1.DEM COP who
'Who is this?'
- b. *jina lako ni nani?*
[G5]name AG5.2SG.POSS COP who
'What (lit.: 'who?') is your name?'
- (36) a. *huyu ni nani?*
AG1.DEM COP who
'[About an unknown animal:] What is this?'
- b. *jina la mnyama huyu ni nani?*
[G5]name AG5.CON G1.animal AG1.DEM COP who
'What (lit.: 'who?') is the name of this animal?'
- c. *ni nani a-li-ye-ku-uma?*
COP who AG1-PST-REL-2SG-bite
'[Someone walks into a room with a swelling on his arm, a wound:]
What bit/stung you? (lit.: 'It is who that bit/stung you?')'
- (37) a. *huu ni nini?*
AG3.DEM COP what
'[About an unknown tree:] What is this?'
- b. *jina la mti huu ni nini?*
[G5]name AG5.CON G3.tree AG3.DEM COP what
'What is the name of this tree?'

Having made this reservation, let us now consider languages that use 'who?' in questions about living things only when it is the name of a folk genus as such that is at stake and not the classification of a living thing. Just as in questions about other kinds of proper names a distinction needs to be made here at least between languages allowing for such questions to be constructed as [Who is X? (by name)], on the one hand, and as [Who is X's name?] and/or [Who is X named?], on the other.

I found only two languages where a question about the name of a folk genus can be constructed with [Who is X?], viz. Jaru/Djaru (Pama-Nyungan; WJa2; Australia; Tsunoda 1995:68) (38) and Yabem/Jabêm (Austronesian, Western

Oceanic, Huon Gulf; Papua New Guinea; Dempwolff 1939:§67 via Bradshaw & Czobor 2005:76-78; Zahn 1940:56) (39).

Jaru

(38) [A:] *ɲana ɲila?*
who that

[B:] *ɲiji, ɲana ɲila*
indeed who that

‘[A man (A) who has never seen a buffalo comes across one, and asks his friend (B):] What is (the name of) that (animal)? [B:] Indeed, (I wonder) what that is’ (Tsunoda 1995:68)

Yabem

(39) a. [A:] *ka tonaj asa?* [B:] *nip*
tree DEM who coconut

‘[A:] What is that tree called? (lit.: ‘That tree is who?’) [B:] A coconut palm’ (Dempwolff 1939:§67 via Bradshaw & Czobor 2005:76-78)

b. [A:] *asa ka kê-kô tonaj?* [B:] *papaja kê-kô*
who tree 3SG-REAL.stand DEM papaya 3SG-REAL.stand
‘[A:] What is the tree there? (lit.: ‘Who, (it is) a tree, stands there?’) [B:] It’s papaya’ (Zahn 1940:56)³⁷

Note that when the question is not about the name of a folk genus as such the

³⁷ Here, both Dempwolff and Zahn, who worked together a lot (cf. Ross 2002:270), would have characterized the word *ka* ‘tree’ as being in apposition to the interrogative *asa*. This can be deduced from the following statement by Dempwolff (1939:§67 via Bradshaw & Czobor 2005:76-78: “If the situation is known to the extent that one is asking about a person belonging to some kind of group, then the noun in question will be placed in apposition after *asa*, as in *aôm asa âwê* [‘you who woman’] ‘which woman are you?’, *aê Samalitô* ‘I am S.’”). Similarly, Zahn (1940:55) notes: “Wenn dem Fragenden bekannt ist, zu welcher Gruppe eine Person gehört, über die er den Eigennamen oder eine bestimmte Angabe erfragen will, so kann *asa* die Bezeichnung dieser Gruppe als substantivische Apposition hinter sich nehmen, z.B. *asa ɲapalê kêtaŋ* ‘wer, ein Kind, weint?’ [which in German would correspond to] ‘welches Kind weint?’, [wir] dürfen aber nicht einen Hinweis [i.e., a demonstrative] als Antwort erwarten”. Alternatively, since genitives precede their heads in Yabem, *asa* could be construed as a possessor of *ka*, so that (39b) would be largely comparable to the Tuvaluan example (31b) above. Although normally the genitive relation between the two nominals is overtly marked in Yabem, “[w]here possessor is non-specific [i]n some cases the possessor simply precedes the possessed in a lexicalised compound” (Ross 2002:280; other modifiers, such as adjectives, quantifiers and demonstratives, are all postposed to nominals in Yabem, Ross 2002:279). Still, the appositional interpretation seems to be preferable. Consider also example (44) below.

interrogative ‘what?’ is used, as can be illustrated with examples in (40) from Yabem.

Yabem

- (40) a. [A:] *gêŋ baliŋ tonec asàgeŋ?* [B:] *moac*
 thing long DEM what snake
 ‘[A:] What is this long thing? [B:] A snake’ (Dempwolff 1939:§46 via Bradshaw & Czobor 2005:38)
- b. [A:] *asàgeŋ kê-taŋ?* [B:] *moc teŋ*
 what 3SG-REAL.cry bird one
 ‘[A:] What’s that sound? (lit.: ‘What cried?’) [B:] Some bird’ (Dempwolff 1939:§67 via Bradshaw & Czobor 2005:76-78)
- c. [A:] *lau sê-niŋ asàgeŋ?* [B:] *sê-niŋ mo*
 people 3PL-IRR.eat what 3PL-IRR.eat taro
 ‘[A:] What will the people eat? [B:] They will eat taro’ (Dempwolff 1939:§46 via Bradshaw & Czobor 2005:38)
- d. [A:] *ŋawê tonan asàgeŋ?* [B:] *mo*
 sprout DEM what taro
 ‘[A:] What is this sprout? [B:] A taro’ (Zahn 1940:53)

The difference between (39a) and (40d), for instance, is particularly illustrative. This difference seems to be due to the fact that *ka* ‘tree’ is a life form and *ŋawê* ‘sprout’ is not. That is, whereas *nip* ‘coconut palm’ is a kind of *ka* ‘tree’, *mo* ‘taro’ is not a kind of *ŋawê* ‘sprout’.³⁸ In other words, in (39a) the question is about identifying a subordinate category with respect to the established superordinate category *ka* ‘tree’ within a transitive hierarchy of categories (“wenn es sich um eine Einzelart aus einer Gattung handelt”, Zahn 1940:56), whereas in (40d) the question is about “reclassifying” the entity from one category, viz. sprouts, in terms of another category, viz. plants (or the like), where the two categories do not form a single hierarchy.

More languages have been found where a question about the name of a folk genus can be constructed [Who is X’s name?] and/or [Who is X named?]. All these languages also make part of the group summarized in (27) in Section II.3.3.3. In particular, this use has been found in four Bantu languages, viz. Tswana, Tswapong, Kgalagadi and Subiya/Kuhane (41), and two Cushitic languages, Hadendowa Bedawi and Libido (42). This use may also be possible in some of the Australian languages mentioned in Section II.3.3.3, but I lack clear data for these languages.

³⁸ A ‘sprout from a seed’ or a ‘sprout from a root’, for instance, would be kinds of ‘sprout’.

Subiya/Kuhane (Niger-Congo, Bantu K50; Namibia & Botswana; Ndana Ndana, p.c.)

- (41) a. *i-zina lye lyi samu njeni?*
 AUG-name of this tree COP.who
 ‘What (lit.: ‘who?’) is the name of this tree?’
- b. *i-lyi samu lyi sumpwa ni?*
 AUG-this tree it.is.called who
 ‘What (lit.: ‘who?’) is this tree called?’

Libido (Afro-Asiatic, Highland East Cushitic; Ethiopia; Joachim Crass, p.c.)

- (42) a. *ka hakk’an summi ’aye?*
 this tree.GEN name.NOM who
 ‘What (lit.: ‘who?’) is the name of this tree?’
- b. *ka hakk’a ’aye yaka’o?*
 this tree.ACC who they.say
 ‘What (lit.: ‘who?’) do they call this tree?’

3.3.6 Proper names of inanimate things

Besides place-names and temporal names, names of many other inanimate things may function as proper names as well, such as names of ships, buildings, works of art, brands, numbers, letters, currencies, languages, colours, diseases, etc. (cf. Van Langendonck 2007:218-246). For some of these names, such as names of ships (e.g., *the Titanic*, *the Queen Mary*) or buildings (e.g., *the Empire State Building* or *the White House*), proper name use may be more typical than for others. In other words, some names of inanimate things may tend to be construed as proprial or at least proprio-appellative lemmas, while others tend to be construed as purely appellative lemmas.

Some names of inanimate things which typically contain people or are directly related to people in some way can be used metonymically to refer to the people they are related to. In questions, this often seems to license the use of ‘who?’, as in (43).

English

- (43) *Who has taken this decision, the White House or the Capitol?*

Obviously, this is neither a non-prototypical combination of values nor a lack of differentiation between ‘who?’ and ‘what?’.³⁹ A direct personification of the names of things without implied metonymy (as can be found with domestic

³⁹ Consider also the discussion in Section III.2.3.4.1.3.4 on “things denoting persons” in Biblical Hebrew and the use of ‘who?’ in questions about such “things”.

animals, cf. Section II.3.3.2) appears to be much less common, though.

I have not found unequivocal examples of questions about proper names of inanimate things that do not use the words ‘name’ or ‘to name, to call’. However, example (44) from Yabem can probably be interpreted this way.

Yabem (Austronesian, Western Oceanic, Huon Gulf; Papua New Guinea)

- (44) [A:] *siasi tonaŋ kê-tê asa wê?*
 Siasi.person DEM 3SG-REAL.perform who dance
 [B:] *eŋ kê-tê wê siac*
 3SG 3SG-REAL.perform dance *sia* (a dance name)
 ‘[A:] What is the dance that Siasi man is dancing? (lit.: ‘Who, (it is) a dance, is that Siasi man performing?’) [B:] He is dancing the dance called *sia*’ (Zahn 1940:56)

As was discussed in the footnote to a structurally similar example (39b) above, the interrogative *asa* in (44) does not seem to function attributively. Rather, the nominal *wê* ‘dance, song’ is used in apposition to *asa*, just as in the answer to (44), the dance name *siac* stands in apposition to the less specific *wê* ‘dance, song’.

A somewhat different situation can be found in the Yankunytjatjara example (45b).

Yankunytjatjara (Pama-Nyungan, Western Desert/Wati; Australia)

- (45) a. *ngana-lu-nta yu-ngu?*
 IPW-ERG-2SG.ACC give-PST
 ‘Who gave it to you?’ (Goddard 1983:8-24.ii via Mushin 1995:19)
 b. *mayu ngana-nya watya-la?*
 tune IPW-ACC tell-IMP
 ‘Tell (us) what tune?’ (Goddard 1983:8-24.iii via Mushin 1995:19)

In (45b), the interrogative *ngana-*, otherwise meaning ‘who?’, as in (45a), is used attributively to the nominal designating a thing to ask about the proper name of this thing. As such, this example is interesting, but the attributive use of interrogatives is outside of the scope of the present study. I do not know whether a non-attributive use of *ngana-* in questions about proper names of things is possible in Yankunytjatjara.

Are there languages that allow (or require) questions about proper names of inanimate things to be constructed as [Who is X’s name?] and/or [Who is X named?]? Although I have not controlled this specifically, my impression is that this must be possible at least in the same four Bantu and two Cushitic languages that allow (or require) the use of ‘who?’ in questions about names of folk genera (cf. Section II.3.3.5), viz. Tswana, Tswapong, Kgalagadi, Subiya/Kuhane (Bantu) and Hadendowa Bedawi and Libido (Cushitic) respectively. Here again, I lack

clear data for the Australian languages mentioned in Section II.3.3.5.

3.3.7 Pure autonyms

Pure autonyms are metalinguistic names, i.e. linguistic expressions that refer to themselves, such as *stand for* and *about* in the phrase *the words ‘stand for’ and ‘about’* (cf. Van Langendonck 2007:246-249). Pure autonymy should, for instance, be distinguished from “reference to conceptual entities, as in *the notion of ‘soul’*” (Van Langendonck 2007:247). In many respects autonyms behave like proper names and should be considered as “a subclass of proper names in their own right” (Van Langendonck 2007:95, 246-249).

As already mentioned in Section II.3.1, strictly speaking, any answer to a question about the name of an entity, such as *What is your name?*, *What is the name of this tree?*, *What do they call this tree?*, would be an autonym. However, here, I am interested in examples such as (46).

Hadendowa Bedawi/Beja (Afro-Asiatic, North Cushitic; Sudan)

- (46) *oo-tam* *ʔaab* *eedna* *t-ʔarabyeet-iib?*
 ART.M.SG.ACC-sorghum.ball who.ACC say.IPFV.3PL ART.F.SG-Arabic-in
 ‘What is *sorghum ball* in Arabic? (lit.: ‘Who do they call *sorghum ball* in Arabic?’)’ (Martine Vanhove & Mohamed-Tahir Hamid Ahmed, p.c.)

Unlike in questions such as *What is the name of this tree?*, in (46) the entity whose proper name is questioned about, viz. *tam*, is used autonymously itself, i.e. it means ‘the word *sorghum ball*’ here.

Besides Hadendowa Bedawi, the only other language in my sample where ‘who?’ may be possible in such questions is Libido, another Cushitic language (47).⁴⁰

Libido (Afro-Asiatic, Highland East Cushitic; Ethiopia; Joachim Crass, p.c.)

- (47) a. *libitt'isan* “sheep”-a ‘aye *yaka’o?*
 Libido.language.in “sheep”-ACC who they.say
 ‘What is *sheep* in Libido? (lit.: ‘Who do they call *sheep* in Libido?’)’

It should be mentioned, though, that according to Joachim Crass (p.c.), not all

⁴⁰ I do not consider here languages like Rapa Nui (Austronesian; Chile), where it seems that only those autonyms can be questioned with ‘who?’ that are autonyms of nominals belonging to the morphosyntactic class of proper names, viz. proper names of persons, some places and months. Cf. example (32) above and the accompanying footnote. Similarly, I would not consider here a language where only those autonyms can be questioned with ‘who?’ that are autonyms of (grammatically animate) nominals denoting living things and where the names of those kinds of entities are questioned with ‘who?’.

speakers accept the use of 'aye 'who?' in (47).

3.4 'Who?'-dominance in NAME-questions: summary

As discussed in Section II.3.2, three major kinds of constructions for questioning proper names may be distinguished: (i) [IPW is X? (by name)], (ii) [IPW is the name of X?] and (iii) [IPW is X named?].⁴¹ For most purposes, it has proved best to treat the latter two constructions together. This grouping is also justified by the fact that only the latter two constructions make explicit use of the notion of naming. In Section II.3.3, I examined seven types of contexts (proper names) where 'who?'-dominance in the case of the non-prototypical combination of values [thing + identification + proper name] has been found in the languages of my sample. The seven types of contexts are as follows: (i) proper names of persons in (Section II.3.3.1), (ii) proper names of domestic animals (Section II.3.3.2), (iii) toponyms (Section II.3.3.3), (iv) temporal proper names (Section II.3.3.4), (v) names of folk genera (Section II.3.3.5), (vi) proper names of inanimate things (Section II.3.3.6), and (vii) pure autonyms (Section II.3.3.7). It appears that these contexts can be organized in an implicational hierarchy as regards the possibility of 'who?'-dominance. In fact, it seems practical to distinguish two subtypes of this hierarchy, viz. one for constructions [Who is the name of X?] and [Who is X named?] (48) and one for construction [Who is X? (by name)] (for non-persons) (49).

- (48) *'who?'-dominance hierarchy for constructions [IPW is the name of X?] and [IPW is X named?]:*

personal proper names < (proper names of domestic animals) < toponyms
< (temporal proper names) < names of folk genera (<) proper names of
inanimate things < pure autonyms

- (49) *'who?'-dominance hierarchy for construction [IPW is X? (by name)] (for non-persons):*

(proper names of domestic animals) < toponyms < (temporal proper
names) < names of folk genera (<) proper names of inanimate things

Due to the small number of examples for the contexts further to the right end of the hierarchy and some gaps in my data, for the time being the two hierarchies should be viewed as first approximations only and tendencies rather than absolute universals.

⁴¹ Recall (Section II.3.2) that further variations on these three types of constructions are possible in given languages. For instance, [IPW is X named?] may also be realized as [IPW do they name/call X?], [IPW does X name/call himself/herself/itself?] and the like.

As can be readily observed, the hierarchy in (49) is in fact a reduced version of the hierarchy in (48). Therefore, I will sometimes refer to the two hierarchies summarily as if they represented one hierarchy. Personal proper names are not included in (49) because there is no non-prototypical combination of values in the case of ‘Who is X?’ where X is a person. Proper names of domestic animals are put in brackets in both (48) and (49) because as discussed in Section II.3.3.2, the presence of this context in the hierarchy appears to be mostly dependent on extralinguistic factors, viz. a given linguistic community must have domestic animals and domestic animals must be important enough to be given proper names. Temporal proper names are put in brackets for a different reason, viz. I have too little data on this kind of context for making cross-linguistic generalizations. Furthermore, at least for (49), it seems that temporal proper names may in fact represent an “offshoot” branching from toponyms. In other words, temporal names may be parallel to, rather than implied by contexts such as names of folk genera and proper names of inanimate things. Finally, it is not clear to me for the moment whether the latter two contexts themselves (in both (48) and (49)) are of different levels, with an implicational relation between them, or of the same level, with no implicational relation between them.⁴²

The languages of my sample proved to be distributed very unevenly on the hierarchies represented in (48) and (49). To begin with, while there are more than a hundred languages taking part in the hierarchy in (48), there are just a few in the hierarchy in (49).⁴³ What is more, the latter languages are all confined to the Oceanic subbranch of the Austronesian family, some Australian (mostly Pama-Nyungan) languages and perhaps some non-Austronesian Papuan languages. In this respect one may wonder whether this distribution does not have something to do with the fact that (i) in both Austronesian and Pama-Nyungan languages, proper names often have a very prominent morphosyntactic status (cf. Section II.3.3.1), and that (ii) in both groups, there is a good deal of languages lacking differentiation between ‘who?’ and ‘what?’ (cf. Sections III.4.2 and III.6).⁴⁴

When we move from left to right on the hierarchy, the number of languages involved reduces drastically, i.e. with a magnitude of several times, with every step to the right. For the moment, only a couple of languages have been found that probably get as far as the right end of the hierarchy, although only for one language, Hadendowa Bedawi, the data appear to be fully waterproof.

⁴² In the latter case, they should be separated by a comma rather than the < symbol.

⁴³ Provided, of course, we do not count languages that use ‘who?’ only in questions about proper names of domestic animals.

⁴⁴ What is more, for the Pama-Nyungan languages, it may be argued that the ‘who?’ interrogative involved, viz. *ngana*, used to be a ‘who?, what?’ interrogative (cf. Section III.6.4.1.1).

Remarkably, both the organization of the hierarchy itself and to a certain extent the distribution of languages on the hierarchy appears to correlate with the prototypicality gradation among the different kinds of proper names advanced by Van Langendonck (2007:254):

[W]e can speak of a gradation in proper names [...], i.e. prototypical to nonprototypical to marginal names without an *ad hoc* proprial lemma. First, there is the most important class consisting of personal names, animal names, place names and a few others; second, there are these categories which are called ‘appellative proper names’ by a few German onomasticians: names of brands, languages, diseases, and so on. We can speak here of proprio-appellative lemmas. What is left is marginal categories without an *ad hoc* proprial lemma, like autonyms [...] or words that appear as proper names in a restricted context [...] personal and place proper names are the prototypical proper names.

However, the very uneven character of the distribution of the languages on the ‘who?’-dominance hierarchy in questions about proper names cannot be explained only in terms of the prototypicality gradation among different kinds of proper names. Rather, I believe that the unevenness of the distribution at issue has to do with the fact that it must be relatively easy for a language to overcome the inanimate status of a proper name as a thing said by taking into account its categorical presuppositional meaning when the latter meaning is ‘person’ (cf. Section II.3.2, Section II.3.3.1), thus assimilating in the context of a question, ‘person’ as a categorical presuppositional meaning to PERSON as a kind of entity. At the same time, the next step involving the assimilation in the context of questions of other prototypical proper names, such as place-names, to personal proper names seems to be much more difficult because the basis for this assimilation, which is the (prototypically) proprial function, is much more abstract than the personal semantics was in the first case. It is worth recalling in this respect that propriality often appears to have a prominent status in grammars of the languages that have taken this step. Finally, a yet further step from prototypical proper names to non-prototypical proper names is complicated by the simple fact that the proprial function, which was the only basis for the step from personal proper names to other prototypical proper names, becomes much weaker, much less evident in the case of non-prototypical proper names. Therefore, we should not be surprised that there are so few languages at the extreme right end of the hierarchy. Similarly, we should not be surprised either that the only language for which the data here is fully unequivocal, viz. Hadendowa Bedawi, appears to have developed this use due to some language-specific concurrence of changes in the system of its interrogatives (cf. Section II.3.3.1 on Cushitic languages) rather than to some particular importance of

propriality in its grammar.

Finally, as far as the areal and genetic distribution of the languages with ‘who?’-dominance in questions about proper names is concerned, when the whole hierarchy is taken into account three major foci may be distinguished on a world-wide scale, viz. (i) Bantu and Cushitic languages in Africa, (ii) Austronesian languages in Asia and the Pacific, and (iii) Pama-Nyungan languages in Australia. The distribution becomes much broader when only the left end of the hierarchy, viz. personal proper names, is taken into consideration (Section II.3.3.1).

4 ‘Who?’-dominance in ANIMATE-questions

This section examines ANIMATE-questions, i.e. questions about animate things, especially when these questions are formulated with ‘who?’, which gives rise to a non-prototypical combination of values of the kind [animate thing + classification (+ common noun)]. Languages may differ as to what non-selective interrogative pronominal they use in questions about animate things (Section II.4.1) and as to what they consider as animate in the context of a question by means of a non-selective interrogative pronominal (Section II.4.2).

4.1 ANIMATE-questions: three strategies

As far as the use of non-selective interrogative pronominals with animate things is concerned, three main strategies appear to be found in the languages of the world: (i) ‘what?’ (Section II.4.1.1), (ii) a special interrogative (Section II.4.1.2), (iii) ‘who?’ (Section II.4.1.3). In addition, all languages appear to allow for avoidance strategies. Presumably the most typical one consists in recurring to a construction with an attributive interrogative of the kind *what (kind of) animal?*.

4.1.1 ‘What?’

Given that by default (non-human) animates, viz. biologically living entities, are things, when it comes to choosing a non-selective interrogative pronominal in questions about such animate things, most languages appear to opt for ‘what?’, as Alawa (1) and Candoshi (2) do, for instance.

Alawa (non-Pama-Nyungan, Maran; Australia; Margaret Sharpe, p.c.)

- (1) *nganjini-rrī galnari nyamba?*
 what-ERG bite.3SG.M.PST 2SG.OBL
 ‘[Someone walks into a room with a swelling on his arm, a wound or whatever. A:] What bit you?’

Candoshi (Isolate; Peru; John C. Tuggy, p.c.)

- (2) *maay shiyaa toshirag?*
 what 2SG.ACC stung
 ‘[Someone walks into a room with a swelling on his arm, a wound or whatever. A:] What stung you?’

4.2.2 A special interrogative

Although by default animate things are things, strictly speaking they fall somewhere in between persons and prototypical things, viz. inanimate things (as

delimited in Section I.2.4). Therefore, many languages tend to avoid, whenever possible, using either ‘what?’ or ‘who?’ in questions about non-human animates, cf. Section I.4.2.3.2. Several languages have been found to distinguish three different kinds of interrogatives, one for persons ‘who?’, one for inanimate things ‘what?’ and one for animate things ‘what? (animate thing)’. Since in the present study I assume that the conceptual space for delimiting the prototypical functions of non-selective interrogative pronominals is basically divided in two parts, viz. ‘who?’ vs. ‘what?’ (cf. Section I.2.6), the use of a special non-selective interrogative pronominal for animate things should be viewed as an avoidance strategy.

A system with a special non-selective interrogative pronominal for animate things can be illustrated with the interrogatives of the Uto-Aztecan language Southern Paiute in (3).

Kaibab variety of Southern Paiute (Northern Uto-Aztecan, Numic; USA; Sapir 1930:208-209)

- (3) *aŋa-* ‘who? (human)’
 in:i- ‘what? (AN)’
 impi- ‘what? (INAN)’

Sapir (1930:208) glosses the animate interrogative *in:i-* as ‘who?, what?, of what sort?’ and characterizes it as “a generalized animate interrogative pronoun, referring to an animal, an animate being not known to be definitely human or animal, or a person of undefined characteristics”. (4) is the only example Sapir provides to illustrate the use of *in:i-* in questions about “a person of undefined characteristics”.

Kaibab variety of Southern Paiute (Sapir 1930:208)

- (4) *in:i-ntcan* *niŋwúruni?*
 IPW-PST.me person.make.me
 ‘Who ever respected me? (lit.: ‘Who ever made me a person?’)’

However, given that (4) seems to be a rhetorical question, *in:i-* here could also be interpreted as having a general animate reference (i.e., not inquiring specifically about a *person* of undefined characteristics) in the sense that not even an animal has ever respected the speaker, let alone a human.

The opposition between human, (non-human) animate and inanimate interrogatives may be further subdivided according to additional parameters. Tucano, for instance, is described by Ramirez (1997:328-332) as having five non-selective interrogative pronominals (5).

Tucano (Eastern Tucanoan; Brazil & Colombia; Ramirez 1997:328-332)

- | | | | |
|-----|---------------|---------|-----------------------------|
| (5) | <i>noá</i> | dōá | ‘who? (human SG or PL)’ |
| | <i>yě’é</i> | yě’é | ‘what? (INAN SG or PL)’ |
| | <i>yamĩ</i> | yābá-gĩ | ‘what? (non-human AN.M.SG)’ |
| | <i>yamó</i> | yābá-go | ‘what? (non-human AN.F.SG)’ |
| | <i>yamârã</i> | yābâ-dã | ‘what? (non-human AN.PL)’ |

Examples (6-8) illustrate the use of some of these interrogatives.

Tucano

- | | | | |
|-----|--|-----------------------------|--------------------------------|
| (6) | <i>noá</i> | <i>a'teré</i> | <i>weerósari?</i> |
| | dōá | a'té-de | weé-do-sa-di |
| | IPW | DEM.INAN-REF | do-INAN-NON<VISUAL>.NON<PST>-Q |
| | ‘[A:] Who will do this? [B: Pedro]’ (Ramirez 1997:329) ⁴⁵ | | |
| (7) | <i>yēeré</i> | <i>iasarí?</i> | |
| | yě’é-de | iá-sa-di | |
| | IPW-REF | want-NON<VISUAL>.NON<PST>-Q | |
| | ‘What do (you) want?’ (Ramirez 1997:331) | | |
| (8) | <i>yamârãre</i> | <i>wēheátí?</i> | |
| | yābâ-dã-de | wēhé-a-ti | |
| | IPW-AN.PL-REF | kill-VISUAL.RECENT.PST-Q | |
| | ‘What (animate things) did (you) kill?’ (Ramirez 1997:331) | | |

The same subdivision into human, animate and inanimate non-selective interrogative pronominals seems to be found in a few other Eastern Tucanoan as well, such as Cubeo (Morse & Maxwell 1999; Ferguson *et al.* 2000), Desano (Miller 1999) and probably Carapana (Metzger 1981).

An Arawakan language, Machiguenga, allows for a similar distinction in a somewhat different way (9).

⁴⁵ Ramirez (1997:330) points out that the use of the inanimate subject agreement marker *-ro* |do| on the verb is somewhat unexpected here and that an animate suffix, viz. *-gĩ* AN.M.SG or *-rã* AN.PL, can be used here instead. Similarly, the inanimate subject agreement marker *-ro* and the animate plural *-rã* can be used interchangeably with the selective interrogative *nohá* |dōhá| ‘which ones? (AN.PL)’. In this respect, note that in origin, *noá* |dōá| itself is a plural form of ‘who?’. Compare, for instance, Desano (Eastern Tucanoan; Brazil & Colombia) *dōã* ‘who? (PL)’ vs. *dĩĩ* ‘who? (M.SG)’ and *digo* ‘who? (F.SG)’ (Miller 1999:32). In Desano, the plural form *dōã* also seems to be used when the sex of the person in question is unknown (‘if the ‘who’ is indefinite”, Miller 1999:32).

Machiguenga (Arawakan; Peru; Snell 1998; Betty Snell, p.c.)

- (9) *tsini* 'who? (human)' (rare)
tyani 'who? (human)'⁴⁶
tata 'what? (non-human AN or INAN)'
tatuita 'what?, what is it? (non-human AN)' (from *tata iita* 'what is it (AN)?')⁴⁷
taita, tatoita 'what?, what is it? (INAN)' (from *tata oita* 'what is it (INAN)?')

A somewhat similar situation can be found in the Algonquian language Blackfoot (10).

Blackfoot (Algonquian, Plains Algonquian; Canada & USA; Frantz 1991:135-136; <http://www.fp.ucalgary.ca/blackfoot/syntax/syntax.htm>)

- (10) *ta(h)káa* 'who?' (and obviative *tsikáa*)
tsá anistápssí-S 'what?, what is it? (non-human AN)' (*anistápssí*- 'be', used with animate subjects)⁴⁸
tsá anistápíí-S, 'what?, what is it?' (INAN) (*anistápíí*- 'be', used with inanimate subjects), also fused without subject marking as *tsáánistápi* or *tsánistapi*⁴⁹
ááhsa 'what?' (INAN)

The use of these Blackfoot interrogatives is illustrated in (11-14).

⁴⁶ The interrogative *tyani* can also mean 'where is he/she/it (AN)?, which one (AN)?', which is its original use. The inanimate counterpart of *tyani*, viz. *tyatí*, appears to be used only selectively as 'which one (INAN)?' or as locative 'where is it (INAN)?' (Betty Snell, p.c.).

⁴⁷ The forms *iita* and *oita* are glossed as *i-i-t-a* 3M-name-EP-NON<FUT>.REFL and *o-i-t-a* 3F-name-EP-NON<FUT>.REFL respectively. In Machiguenga, the verb *-i-* 'name, call' is found only in combination with *tata* 'what?'. However, it can still be found in regular use in other closely related Arawakan languages, such as Apurucayali Asheninca (cf. Section III.8.2, especially Section III.8.2.1). The association between the masculine *iita* and the animate meaning and between the feminine *oita* and the inanimate meaning is due to the fact that in Machiguenga, which has a so-called combined gender system (cf. Section III.8.2.3), all masculine nominals are also animate, whereas all inanimates are also feminine (cf. Snell 1998:28).

⁴⁸ In combination with other verbs, the interrogative *tsá* can be used to mean 'why?', 'when?', 'how many?', '(do or say) how?' (cf. Frantz 1991:137-139).

⁴⁹ I do not know whether there are similar fused animate forms without subject marking, such as **tsáánistápssí* or **tsánistapssi*.

Blackfoot

- (11) *ta(h)káa áwaasai'níwa?*
 |ta(h)káa á-waasai'ni-wa|
 IPW DUR-cry-3SG
 'Who is crying?' (Frantz 1991:135)
- (12) *tsá anistápssíwa ánniksi áyo'kaiksi?*
 |tsá anistapssi-wa ann-iksi á-yo'kaa-iksi|
 IPW be[S.AN]-3.NON<AFFIRMATIVE>DEM-3PL DUR-sleep[S.AN]-3PL
 'What (AN) are they that are sleeping?' (Frantz 1991:136)
- (13) a. *tsá anistápssíwa annáhka*
 |tsá anistapssi-wa ann-wa-hka|
 IPW be[S.AN]-3.NON<AFFIRMATIVE>DEM-3SG-INVISIBLE
kitohpómmatawahka?
 |kit-ohpommat-a:-wa-hka|
 2-buy[OBJ.AN]-DIRECT-3SG-INVISIBLE
 'What (AN) did you buy? (lit.: 'What is it (AN) that you bought?')'
 (Frantz 1991:136)
- b. *tsá anistápiiwa anníhka*
 |tsá anistapii-wa ann-yi-hka|
 IPW be[S.INAN]-3.NON<AFFIRMATIVE> DEM-INAN.SG-INVISIBLE
kitohpómmatoohpihka?
 |kit-ohpommatooh-pi-hka|
 2-buy[OBJ.INAN]-NMLZ-INAN.SG-INVISIBLE
 'What (INAN) did you buy? (lit.: 'What is it (AN) that you bought?')'
 (Frantz 1991:136)
- c. *ááhsa anníhka kitohpómmatoohpihka?*
 |ááhsa ann-yi-hka kit-ohpommatooh-pi-hka|
 IPW DEM-INAN.SG-INVISIBLE 2-buy[OBJ.INAN]-NMLZ-INAN.SG-INVISIBLE
 'What (INAN) did you buy? (lit.: 'What is it (AN) that you bought?')'
 (Frantz 1991:136)
- (14) *tsánistapi nóóyiwaats?*
 |tsaanistapi n-ooyi-waatsiks|
 IPW AFF-eat[S.AN(vi)]-3SG.NON<AFFIRMATIVE>
 'He ate what?' (<http://www.fp.ucalgary.ca/blackfoot/syntax/syntax.htm>)

Cross-linguistically, the three-way distinction between human, (non-human) animate and inanimate interrogatives as presented in this section appears to be very rare. In all probability, this has to do with the following reasons. To begin with, semantically, such a three-way system is obviously more detailed than the usual two-way system distinguishing just persons and non-persons.

Consequently, it commits the speaker to consciously making more detailed decisions on the kind of entity involved than is strictly speaking desirable in the context of a question (which implies some lack of knowledge) in view of the fact that the most fundamental form of human categorization of entities is the distinction between persons and non-persons. Of course, decisions may be much less conscious, for instance when interrogative pronominals distinguish gender. Gender is a grammatical category, which implies that to a large extent, decisions are imposed on the speaker by the gender system of the language. Note in this respect that in all the cases of a three-way distinction between human, (non-human) animate and inanimate interrogatives cited above the respective languages distinguish animate and inanimate genders and this gender distinction is expressed in the non-selective interrogative pronominals. As such, the opposition between an animate and inanimate gender is already rather uncommon cross-linguistically (cf. Corbett 2005). Furthermore, in systems with an animate and an inanimate gender, there seems to be a tendency to treat humans as animates, i.e. to go without a separate exclusively human gender. Another problem is that even in languages with genders, it is not uncommon for non-selective interrogative pronominals to show little sensitivity to gender. Thus, they would often carry no overt gender marking even when other kinds of pronominals do. Non-selective interrogative pronominals also regularly take some default agreement pattern prescribed by the gender system of the language with human or non-human nominals respectively rather than the pattern one may have expected based on the semantics of genders and entities questioned about. All in all, it seems that too many problems have to be overcome for a language to develop a three-way distinction between human, (non-human) animate and inanimate non-selective interrogative pronominals.

4.1.3 'Who?'

Since strictly speaking, animate things fall somewhere in between persons and prototypical things, languages may also opt for treating them as similar to humans, rather than inanimate things or a special case on their own. Indeed, a group of languages has been found to treat non-human animates similarly to humans as regards the use of non-selective interrogative pronominals. Consider, for instance, Russian *kto* or Estonian *kes*, which can be used in both questions about persons (15, 17, 18) and (some) animates (16, 18).

Russian

- (15) *Kto otkryl Ameriku?*
 who discovered America
 'Who discovered America?'

- (16) *Kto eto tebja ukusil? Osa?*
 who this you bit wasp
 '[Looking at a swelling on someone's hand clearly caused by an insect bite:] What stung you? A wasp?'

Estonian (Uralic, Finnic; Estonia)

- (17) *kes täna puudu-b?*
 who today wanting-3SG.PRS
 'Who is absent today?' (<http://dict.ibs.ee>, under *who*)
- (18) *kes sind hammusta-s?*
 who 2SG-PARTITIVE bite-3SG.PST
 '[Someone walks into a room with a swelling on his arm, a wound:] What bit you?' (an animal, not an insect) or 'Who bit you?' (a person) (Petar Kehayov, p.c.)

Consider also examples from Trió/Tiriyó (19) and Ket (20).

Trió/Tiriyó (Northern Carib; Suriname & Brazil; Sérgio Meira, p.c.)

- (19) a. *aki mēē?*
 IPW.AN DEM
 'Who is this? (a person)' or 'What is this? (e.g., an animal)'
- b. *aki ě-eka?*
 IPW.AN 2OBJ-bite.PST
 '[Someone walks into a room with a swelling on his arm, a wound. A:] What bit you?'

Southern Ket (Yeniseian; Russia; Werner 1997:363)

- (20) [A:] *tu-r' bits'e?* [B:] *tu-r' ap qi'p*
 DEM-AN.M IPW.AN.M DEM-AN.M 1SG.GEN grandfather[AN.M]
 [A:] *a tu-r' bits'e?* [B:] *tu-r' da-s'el'*
 and DEM-AN.M IPW.AN.M DEM-AN.M 3SG.AN.M.GEN-reindeer[AN.M]
 '[A:] Who is this? [B:] It's my grandfather. [A:] And what is this? [B:] It's his reindeer'

As discussed in Section I.4.2.3.2, I prefer to consider this and similar uses of the otherwise human non-selective interrogative pronominals [person + identification (+ proper name)] in questions about animate things [animate thing + classification (+ common noun)] as a non-prototypical combination of values rather than a lack of differentiation. The reason is that although when questioned with 'who?' animate things are assimilated to persons, 'who?' here is used (non-prototypically) for [classification (+ common noun)] and not for [identification (+ proper name)] (obviously, because typically animate things do not have proper names), which implies a non-prototypical combination of values [animate thing

⊆ person + classification (+ common noun)].

In what follows, let us consider languages like Russian, Estonian, Tiriyó and Ket in more detail. To begin with, note that there is an important difference between Russian and especially Estonian, on the one hand, and Tiriyó and Ket, on the other. In particular, Ket and Tiriyó have gender systems distinguishing animate and inanimate genders,⁵⁰ whereas Estonian does not have gender at all and Russian has a sex-based gender system, distinguishing masculine, feminine and neuter genders, with the opposition animate vs. inanimate only having the status of so-called “subgenders”.⁵¹ (This is why I have glossed Russian *kto* and Estonian *kes* as ‘who?’ and Tiriyó *aki* as IPW.AN, i.e. “animate interrogative pro-word”, and Ket *bits’e* as IPW.AN.M, i.e. “masculine animate interrogative pro-word”). In this respect, recall that no languages like Russian or Estonian have been found among languages with a three-way distinction between human, (non-human) animate and inanimate interrogatives discussed in Section II.4.1.2. There also appear to be considerably more languages using the same interrogative both in questions about humans and animate things than languages using a special interrogative for animate things, distinct from both ‘who?’ for humans and ‘what?’ for inanimate things. Still, most languages that use the same interrogative both in questions about humans and animate things have gender systems distinguishing animate and inanimate (sub)genders (sometimes, the gender system may be sex-based at the same time), as in Tiriyó, Ket and Russian. Note that the opposite is not true. That is, languages with animacy-based gender systems do not always use the same interrogative both in questions about humans and animate things. As discussed in Section II.4.1.2, some of them may have a separate non-human animate interrogative ‘what?’ next to a human ‘who?’ and an inanimate ‘what?’, while some may simply have a human ‘who?’ and a non-human ‘what?’, as seems to be the case in Arapaho, Cheyenne (both languages are Algic, Plains Algonquian; USA; cf. Section III.7.2), and Yuchi (Isolate; Oklahoma, USA; Linn 2000:471-472),⁵² for instance. Finally, as will be

⁵⁰ In Ket, the animate gender is further subdivided into masculine and feminine animate genders.

⁵¹ A subgender is, as it were, a gender in embryo. The notion of subgender in general and the situation in Russian in particular are discussed by Corbett (1991:161-168). Corbett (1991:163) defines subgenders as “agreement classes which control minimally different sets of agreement, that is, agreements differing for at most a small proportion of the morphosyntactic forms of any of the agreement targets”. A subgender also involves a “dependent target gender”, which is “a target gender consisting of a set of morphological realizations which mark agreement with members of a given agreement class by an opposition involving only syncretism (and no independent forms)” (Corbett 1991:164).

⁵² Yuchi, as described by Linn (2000), distinguishes ten grammatical genders, of which six are

discussed in Section II.4.2, even if in a language with grammaticalized animacy many grammatical animates are indeed questioned with the same interrogative as humans, this is not necessarily possible for all grammatical animates (let alone grammatically inanimate animate things).

Languages like Estonian that do not have animacy-based gender but use ‘who?’ in questions about animates appear to be very rare. What is more, it seems that typically they are spoken in the vicinity of some languages that have a gender system which is (at least, to an important extent) based on animacy or their ancestor languages used to have such gender systems themselves. Thus, no gender system is reconstructed for any ancestor language of Estonian, but Estonian is spoken in the immediate vicinity of Russian and for almost three centuries Estonia was part of Russia and (later) USSR.

Nganasan is another Uralic language that appears to be somewhat similar to Estonian. It is spoken in the Russian Far North and most Nganasans are bilingual in Russian. Like Estonian, Nganasan does not have gender and no gender system is reconstructed for any ancestor language of Nganasan. Nganasan has two non-selective interrogative pronominals *sili* ‘who?’ and *maa* ‘what?’. Although animate things are normally questioned with *maa* ‘what?’, there appear to be some contexts where only *sili* ‘who?’ is possible, as in (21), where an animate thing is the agent of an action of which a human is the patient.

Avam Nganasan (Uralic, Samoyedic; Russia; Valentin Goussev, p.c.)

- (21) *sili*/**maa* *tənə* *sakələ-hu* *d’üdü-mtə?*
 who/*what 2SG bite-Q.PST arm-ACC.2SG
 ‘[Someone walks into a room with a swelling on his arm, a wound. A:]
 What bit you in your arm?’

Another language that may be similar to Estonian is Hadza, a linguistic isolate spoken in Tanzania. Hadza has an interrogative *akwadza* ‘who?’⁵³ (22)

for Yuchi persons in varying kin relationships to the speaker (with “a few symbolic exceptions to this that occur in traditional tales and in some references to God”, Linn 2000:371), three for inanimate objects (also some biologically animate things, such as plants), and one for animals, all non-Yuchis and a few inanimate things (cf. also Mithun 1999:103). At the same time, among the non-selective interrogative pronominals, the only distinction is claimed by Linn (2000:471) to be between the human interrogative *wanō* ‘who?’ and non-human *wikæ* ‘what?’, even though *wa-nō* ‘who?’ is composed of the interrogative root *wa-* and the “animate classifier” -*nō* (“animate” here implies humans, both Yuchi and non-Yuchi, and other grammatically animate nominals).

⁵³ This interrogative may go back to a combination [*’aku* ‘what (kind of) [N]?, which [N]?’ + *hadza* ‘Hadza, human being’], i.e. ‘what/which person?’. There is also another interrogative ‘who?’ in Hadza, viz. *tame* ‘who?’, which seems to go back to a similar combination, viz. [**ta-*

that can also be used in questions about animate things (23).

Hadza (Isolate; Tanzania)

(22) *'akwadza-ne'e?*

who-1SG.M

'Who am I? (pronounced by a man)' (Bonny Sands, p.c.)

(23) [A:] *'akwadza-ko hako same-kwa-ta 'ono [...]*

who-SG.F DEM.SG.F fed.on-1SG-3SG.F 1SG.M

'akanabe-na 'akwadza 'itita? [B:] //uhi-yako hako

be.called-2SG who 2SG mosquito-3SG.F.COP DEM.SG.F

'[A:] What is this thing that has eaten me up? [...] What do you call it?

[B:] Mosquito' (Bala 1998, via Bonny Sands, p.c.)⁵⁴

Unlike Estonian, Hadza does have gender, viz. masculine vs. feminine, but gender in Hadza is based on sex, not animacy. However, Hadza is surrounded by northeastern Bantu languages many of which have expanded their human agreement pattern (itself originally a functional extension of the agreement pattern of the so-called gender 1 that contains mostly human nouns) as the animate agreement pattern (cf. Van de Velde 2006).⁵⁵ In many of these Bantu languages, it seems to be possible to use 'who?' in questions about animals, as illustrated with (24) from Swahili.

Swahili (Niger-Congo, Bantu G40; Tanzania, Kenya, etc.)

(24) a. *nani a-li-ku-saidia?*

who AG1-PST-2SG-help

'Who helped you?' (<http://www.yale.edu/swahili>)

b. *ni nani a-li-ye-ku-uma?*

COP who AG1-PST-REL-2SG-bite

'[Someone walks into a room with a swelling on his arm, a wound:]
What bit/stung you? (lit.: 'It is who that bit/stung you?')' (Assibi Amidu, p.c.)

While Hadza is comparable to Estonian, Kathmandu Newar, a Sino-Tibetan

'how?' + *leme* 'man'].

⁵⁴ I am grateful to Bonny Sands for making Bala (1998) and several other Hadza texts available to me and for an interesting discussion of the Hadza data. For convenience sake, I have slightly modified the glosses she provided.

⁵⁵ According to a different analysis (e.g., Corbett 1991:43-48, 248-259), implying among other things the acceptance of the concept of the so-called hybrid nouns (cf. Corbett 1991:183-184), these Bantu languages have (partly) reassigned most animates to the human gender 1, which therefore has become the animate gender.

language spoken in Nepal, may be compared to Russian, as a language with an emergent animacy-based gender that also uses the same interrogative in both questions about persons (25, 26, 27a) and (some) animates (26, 27b).

Kathmandu Newar (Sino-Tibetan, Tibeto-Burman; Nepal)

- (25) *su wɔl-ɔ?*
 who come-PFV.DISJUNCT
 ‘Who has come?’ (Hargreaves 2003:382)
- (26) *thwa su kha:?*
 DEM who COP.STATIVE
 ‘Who is this? (a person)’ or ‘What is this? (an animal)’ (Kazuyuki Kiryu, p.c.)
- (27) a. *chanta su-nā: vā:-nya:ta?*
 2SG.DAT who-ERG teeth-bite.PST.DISJUNCT
 ‘[Someone walks into a room with a swelling on his arm, a wound and there are clear traces of human teeth. A:] Who bit you?’ (Kazuyuki Kiryu, p.c.)
- b. *chanta su-nā: nya:ta?*
 2SG.DAT who-ERG bite.PST.DISJUNCT
 ‘[Someone walks into a room with a swelling on his arm, a wound presumably caused by an animal bite. A:] What bit you?’ (Kazuyuki Kiryu, p.c.)

As far as I can judge from the data available to me, Kathmandu Newar seems to be an exception both among the related Sino-Tibetan languages and the neighbouring Indo-European languages. Similarly to other Sino-Tibetan languages, Kathmandu Newar is also usually described as lacking gender. However, Kathmandu Newar has an extensive system of numeral classifiers (more than 50 in total), which includes two “general classifiers”, the animate classifier *-mhɔ/-hmɔ* and the inanimate *-gu(l)* (cf. Hargreaves 2003:374-375, 378; Bhaskararao & Joshi 1985 via Aikhenvald 2000:212-213). Note that *animate* here includes humans. What is more, according to Hargreaves (2003:378-379), in Kathmandu Newar, attributively used demonstratives and genitive nouns and pronouns “may be optionally marked with the attributive suffixes [...]: *-gu* ‘inanimate head noun’; *-mhɔ* ‘animate head noun’” and adjectives (as well as verbs of relative clauses) “are marked with the ‘attributive’ [...] suffixes relative to the head noun: *-gu* ‘inanimate’, *-mhɔ* ‘animate’, and *-pĩ:* ‘animate plural’”. In other words, we may probably speak of an emergent animacy-based gender system in Kathmandu Newar. In this respect, Kathmandu Newar may be viewed as being similar to Russian, with the difference that Russian already has a full-fledged sex-based gender system on the basis of which the animacy-based

subgenders have evolved.

Lezgi, a Northeast Caucasian language spoken in Russia and Azerbaijan, can be cited here as the only language I know of that (i) does not have gender but (ii) allows for the same interrogative to be used in both questions about persons (28a) and animates (28b), *and* (iii) where this use appears to be due to the fact that its ancestor language used to have a gender system largely based on animacy (although Russian influence cannot be completely excluded either).

Lezgi (Northeast Caucasian, Lezgi; Russia & Azerbaijan)

- (28) a. *kwez ni lahana? Ahmed-a, Kerim-a?*
 2PL.DAT who.ERG say.PST PROP-ERG PROP-ERG
 ‘Who has told you? Ahmed? Kerim?’ (Akhmedov 1997:115)
- b. *wun ni k’asna?*
 2SG.ABS who.ERG bite.PST
 ‘What bit you? (a mammal, not a snake or insect)’ or ‘Who bit you? (a person)’ (based on Shejxov 2004:84)⁵⁶

Although Lezgi, together with Aghul/Agul and Udi, does not have gender, the remaining six Lezgi languages do (cf. Alekseev 1985:125-133, 2001e:371). Typically, they distinguish four genders: G1 masculine human, G2 feminine human, G3 containing animates and some inanimates, and G4 mostly confined to inanimates. The same four-gender system is also reconstructed for Proto Northeast Caucasian (cf. Table III.3.2:10). That is, Lezgi must have lost gender and the gender system of its ancestor language was largely based on animacy.

It also seems reasonable to assume that the destruction of the gender system in pre-Lezgi has passed through a stage where the opposition was between an animate and an inanimate gender. In this respect, consider, for instance, the situation in the Luchek variety of Rutul, a Lezgi language which generally has preserved the four-gender system. According to Alekseev (1985:90-91), whereas in the singular most verbs distinguish four genders, in the plural, depending on the verb, the distinction may be (i) between human (i.e., with subjects of G1 or G2) vs. non-human (i.e., with subjects of G3 or G4), (ii) between animate (i.e., with subjects of G1 or G2 and animate subjects of G3 or G4) vs. inanimate (i.e.,

⁵⁶ This is a constructed example based on the following discussion in Shejxov (2004:84): “In Lezgi, the pronoun *wuž* ‘who?’ asks about a person or an animal, i.e. about an animate referent, and the pronoun *wuč^h* ‘what?’ about both animate and inanimate referents [i.e., non-humans] [...] the pronoun *wuč^h* ‘what?’ is used [...] when the question is about the agent of an action (mostly with respect to insects, reptiles, etc., but definitely not with respect to bigger representatives of the animal kingdom or persons), e.g.: *wu[n] kü k’asna?* ‘What bit/stung you?’ (provided it is clear that the bite is by an insect or snake, etc., but not a dog, for instance)”.

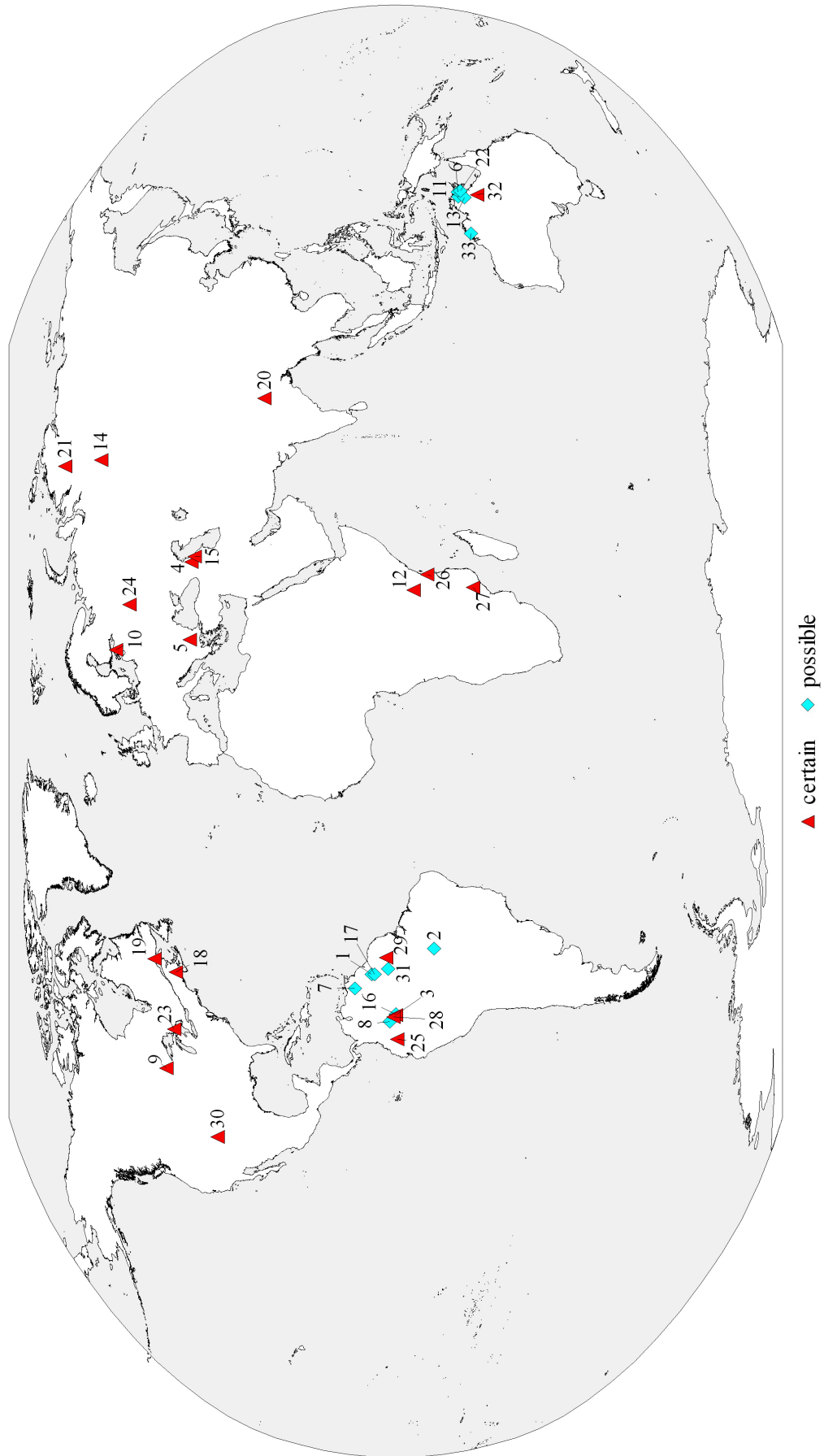
with inanimate subjects of G3 or G4), (iii) between human (i.e., with subjects of G1 or G2) vs. animate (i.e., with animate subjects of G3 or G4) vs. inanimate (with inanimate subjects of G3 or G4), or (iv) animate non-human (with animate subjects of G3 or G4) vs. the rest (i.e., with subjects of G1 or G2 and inanimate subjects of G3 or G4). Outside of the Lezgetic group, we may also cite Botlikh, a language belonging to the Avar-Andic branch of Northeast Caucasian. According to Gudava (1967a) and Magomedbekova (2001), in the singular, Botlikh distinguishes M.HUM (G1), F.HUM (G2) and NON<HUM> (G3) genders. However, on some targets, such as the negation marker *-hi* or question marker *-ma*, the distinction is between animate vs. inanimate. In the plural, the distinction is only between animate vs. inanimate on all targets that can agree. The same system is found in the interrogative pronominals of Botlikh, cf. Table 1. Note also that in the Botlikh interrogative pronominals, gender is neutralized in all cases except absolutive.

Table 1. The interrogative pronominals of Botlikh (Northeast Caucasian, Avar-Andic; Russia; based on Gudava 1967a:301 and Magomedbekova 2001:233)

	SG			PL	
	M.HUM	F.HUM	NON<HUM>	AN	INAN
ABS	<i>ẽ-w/e-w</i>	<i>ẽ/e-j</i>	<i>e-b</i>	<i>ẽo-l/eo-l</i>	<i>e-b</i>
ERG		<i>ɬ:e-di</i>		<i>ɬ:un-di</i>	
GEN		<i>ɬ:e-b</i>		<i>ɬ:un-ɦi</i>	
DAT		<i>ɬ:e-j</i>		<i>ɬ:u-j</i>	

By way of conclusion, consider Map 2 illustrating the distribution of languages that allow the same non-selective interrogative pronominal to be used for both [person + identification (+ proper name)] and [animate thing + classification (+ common noun)]. As can be observed, there are about twenty clear cases of such languages in my sample and about ten more possible candidates. Of course, I may have overlooked some languages of this type because this kind of data is rarely explicitly mentioned in most descriptions. It seems reasonable to expect that the chances to find new languages of this type would be highest for the families or regions which are already known to have at least one such language. It should be kept in mind, though, that even closely related languages may differ as to whether they allow the same non-selective interrogative pronominal to be used for both [person + identification (+ proper name)] and [animate thing + classification (+ common noun)]. For instance, this use is possible in Russian but not in Polish, even though the two are rather closely related Slavic languages and both have subgenders based on animacy.

Map 2. The distribution of languages allowing ‘who?’ in ANIMATE-questions



Language	Family (Ethnologue)
1 Akawaio/Ingarikó*	Carib
2 Apalai*	Carib
3 Barasana	Tucanoan
4 Botlikh	North Caucasian
5 Bulgarian	Indo-European
6 Burarra*	Australian
7 Carib/Galibi*	Carib
8 Carijona*	Carib
9 Chippewa (Southwestern Ojibwa, Ponemah)	Algic
10 Estonian	Uralic
11 Guragone/Gurr-goni*	Australian
12 Hadza/Bali	Khoisan
13 Jawoyn/Djauan*	Australian
14 Ket	Yenisei
15 Lezgian	North Caucasian
16 Macuna*	Tucanoan

Language	Family (Ethnologue)
17 Macushi*	Carib
18 Malecite-Passamaquoddy	Algic
19 Montagnais	Algic
20 Newar	Sino-Tibetan
21 Ngasan (Ust'-Avam)	Uralic
22 Ngandi*	Australian
23 Ottawa/Nishnaabemwin	Algic
24 Russian	Indo-European
25 Secoya	Tucanoan
26 Swahili	Niger-Congo
27 Takwane	Niger-Congo
28 Tatuyo	Tucanoan
29 Trió	Carib
30 Ute (Kapúta & Moǵwáǵi)	Uto-Aztecan
31 Waiwai/Wai Wai*	Carib
32 Wambaya	Australian
33 Worora/Worrorra*	Australian

* possible

As to the genetic distribution, particularly important numbers of the languages of the type at issue have been found in the Algonquian branch of the Algic family in North America, the Carib and Tucanoan families in South America, quite a few (mainly Eastern) Bantu languages of the Niger-Congo phylum in Africa, several Slavic languages in Europe, as well as presumably some non-Pama-Nyungan languages of northern Australia and Yeniseian languages of northern Eurasia, many of which are long-extinct by now. As discussed above, there appears to be a clear correlation between the distribution of the languages that allow the same non-selective interrogative pronominal to be used for both [person + identification (+ proper name)] and [animate thing + classification (+ common noun)] and the distribution of languages that have gender systems largely based on animacy.

4.2 Animacy: cut-off points

Languages allowing for the same non-selective interrogative pronominal to be used for both [person + identification (+ proper name)] and [animate thing + classification (+ common noun)] may differ as to what they categorize as animate (i) in their grammatical system in general (Section II.4.2.1), and (ii) in the context of a question by means of a non-selective interrogative pronominal in particular (Section II.4.2.2).

4.2.1 Cut-off points: in general

The issue of the possible cut-off points on the animacy hierarchy in the grammatical systems of the languages of the world is clearly too complex to be discussed here in detail. Therefore, I will confine myself here to some general remarks.

To begin with, recall that as discussed in Section I.2.4, animacy as a linguistic concept does not need to, and in fact, normally does not coincide with biological animacy. Basically, linguistic (or grammatical) animacy reflects the extent to which a given linguistic system conceives the entity denoted by a nominal as similar to humans and ultimately to SELF. Similarity is a relative concept, which allows for freedom. Moreover, similarity can also be in form rather than in meaning. Consequently, some nominals denoting biologically inanimate things may happen to be characterized as animates, while some nominals denoting humans may happen to be characterized as inanimates. In some border cases, speakers may also vary in their judgements on the grammatical animacy of a given nominal.

Cross-linguistically, nominals denoting certain kinds of entities appear to be more prone to become characterized as grammatically animate than others. Within biologically animate entities, the major divide seems to lie between fauna

and flora. Thus, examples of languages more or less consistently classifying flora entities as animate appear to be hard to find outside of the Algonquian branch of the Algic family in North America and Yeniseian languages of northern Eurasia.⁵⁷ As to fauna entities, larger mammals are almost always classified as grammatically animate. This is less obvious with birds, fish and especially insects, let alone microorganisms.

It is more difficult to make generalizations about biologically inanimate entities that happen to be classified as grammatically animate without running the risk of recurring to semantic criteria that cannot be applied in a principled way, such as *vitality*, *power* and the like. Often, the classification of a given inanimate entity as grammatically animate can be explained by appealing to the mythology of a particular linguistic community. Thus, mythology often seems to be responsible (directly or indirectly) for the fact that cross-linguistically, the names of heavenly bodies are one of the most common groups of inanimate entities characterized as grammatically animate. Metaphor and metonymy are also frequent mechanisms leading to the classification of certain inanimate things as grammatically animate ones. Finally, it should be mentioned that animacy may also be assigned on purely formal reasons or because the nominal used to denote an inanimate thing also (in origin) denotes an animate thing.⁵⁸

4.2.2 Cut-off points: in questions

In languages allowing for the same non-selective interrogative pronominal to be used for both [person + identification (+ proper name)] and [animate thing + classification (+ common noun)], grammatical animacy in the context of a question by means of a non-selective interrogative pronominal is often manifested in a rather different way than it does elsewhere in the grammatical system. What is more, my impression is that *different* here normally means *more restricted*, as well as *less fixed*. That is, (i) not all entities denoted by nominals that are grammatically animate can be questioned with ‘who?’ and (ii) one and the same grammatically animate nominal (in a given meaning) may be questioned with ‘who?’ in one context but with ‘what?’ in another, whereas if (in a given meaning) a nominal is characterized as grammatically animate, it

⁵⁷ Northeast Caucasian languages should probably be included here as well.

⁵⁸ In Russian, for instance *Orjol* ‘Aquila, Vulture (constellation)’ is grammatically animate because *orjol* ‘eagle’ is grammatically animate. Note, however, that at the same time, *Moskvič* ‘Moskvich (car brand)’ is inanimate (although according to Shvedova *et al.* 1980:§1130, it may be animate for some speakers), whereas *moskvič* ‘Muscovite (inhabitant of Moscow)’ is animate. Similarly, *Neptun* ‘Neptune (planet)’ is inanimate, whereas *Neptun* ‘Neptune (god)’ is animate.

normally remains grammatically animate irrespective of the context. In what follows, I will discuss a few examples of what *more restricted* and *less fixed* may imply in particular languages.

To begin with, as was already pointed in Section II.4.1.3, some languages with animacy-based gender systems may simply have a human ‘who?’ and a non-human ‘what?’, as it seems to be the case in Arapaho, Cheyenne (both languages are Algonquian, Plains Algonquian; USA; cf. Section III.7.2), and Yuchi (Isolate; USA), for instance, or they may have a separate non-human animate interrogative ‘what?’ next to a human ‘who?’ and an inanimate ‘what?’ (cf. Section II.4.1.2). It is also possible that a language distinguishing grammatically animate vs. inanimate nominals may use ‘who?’ only in questions about a subset of grammatical animates. Southwestern Ojibwa (Algonquian, Central Algonquian), as spoken in Ponemah, Minnesota, for instance, classifies nominals into an animate and an inanimate gender. Similarly, it distinguishes between the animate interrogative pronominal *awenen* and the inanimate one *awekonen*. However, according to Schwartz & Dunnigan (1986:304), the animate interrogative *awenen* is rather restricted to questions about humans and “large animal[s], such as a bear or a horse”, as in (29a, b), while the inanimate interrogative *awekonen* would be used for insects for instance, as in (29c), even though e.g. *sakime* ‘mosquito’ belongs to the animate gender (cf. also Section III.7.2).

Southwestern Ojibwa of Ponemah, Minnesota (Algonquian, Central Algonquian, Ojibwa; USA; Schwartz & Dunnigan 1986:304)

- (29) a. *awenen kaa-takkwamaat?*
 IPW.AN PST-bite.DIRECT.3OBJ
 ‘Who (person) bit him/them (person)?’
- b. *awenen-an kaa-takkwamikot?*
 IPW.AN-OBV PST-bite.INVERSE.3OBJ
 ‘What/who (“a large animal, such as a bear or a horse”, rather than a person) bit him/them (person)?’
- c. *awekonen kaa-takkwamikot?*
 IPW.INAN PST-bite.INVERSE.3OBJ
 ‘What (e.g., an insect) bit him/them (person)?’

In some languages, the situation may be much more complex than just a simple subdivision of grammatical animates in those that are questioned with ‘who?’ and those that are questioned with ‘what?’. A good example of such a language is provided by Russian.⁵⁹ Roughly speaking, in Russian, all nominals denoting fauna entities (except collectives, cf. Section I.4.2.3.3.2) are

⁵⁹ The situation in other languages may be just as complex as it is in Russian, but I do not have the relevant data.

grammatically animate, whereas all nominals denoting flora entities are inanimate. Furthermore, there are also some grammatically animate nominals denoting biologically inanimate entities, such as *valet* ‘jack (playing card)’, *kozyr* ‘trump’ (cf. Shvedova *et al.* 1980:§1130). Russian has two non-selective interrogative pronominals, *kto* ‘who?’ and *čto* ‘what?’. As a rule of thumb, *kto* is used for persons and fauna entities, as was illustrated in (15-16) above, and *čto* is used for the rest. However, in practice, there are many complications.⁶⁰

To begin with, as discussed in Section I.4.2.3.2, sometimes neither *kto* nor *čto* can be used in a question about a fauna entity. At the same time, in many cases, both *kto* and *čto* can be used in questions about animals. The choice between *kto* and *čto* may then depend on various parameters and often it is difficult to predict which parameter will be given preference. An important parameter, for instance, is the size of the animal: the larger the animal the more likely it is that *kto* will be used rather than *čto*. Indirectly related to the latter parameter is the tendency to use *kto* with mammals and *čto* rather than *kto* with birds, reptiles and especially fish and insects. Furthermore, *čto* rather than *kto* may be preferred if the animal is perceived as unpleasant, repulsive, etc., and on the contrary, *kto* may be preferred if the animal is nice or pleasant in some way. Similarly, if the animal is perceived as food *čto* may be preferred to *kto*. Consider, for instance, examples (30-31).

Russian

- (30) a. *Čto ty segodnja pojmal? Okunja? Ščuku?*
 what.ACC you today caught perch.ACC pike.ACC
 ‘[To someone coming back from fishing:] What have you caught today? A perch? A pike?’ (based on Barulin 1980:36)
- b. *Čto ty segodnja podstrelil? Krolika? Utku?*
 what.ACC you today shoot.down rabbit.ACC duck.ACC
 ‘[To someone coming back from hunting:] What have you shot today? A rabbit? A duck?’
- (31) a. *Čto/kto tam letit? Lebed’ ili gus’?*
 what/who there flies swan or goose
 ‘What is it flying there? A swan or a goose?’ (Houtzagers 2003:203)
- b. *Kogo eto Vy prodajote? [Kakaja zabavanja ptička!]*
 who.ACC this you sell
 ‘[At a pet market:] What are you selling here? [What a funny bird!]

Furthermore, one and the same kind of grammatical animates may be normally questioned with *čto* in one context but with *kto* in another. Thus, insects are

⁶⁰ Cf. Section I.4.2.3.3.2 for complications with collectives.

typically questioned with *čto*, as in (32). However, when insects act on humans, *kto* will normally be used, as in (33).⁶¹

Russian

- (32) a. *Čto eto po stene polžet? Pčela ili osa?*
 what this on wall crawls bee or wasp
 ‘What is it crawling on the wall? A bee or a wasp?’ (based on Barulin 1980:36)
- b. *Čto tam u tebj na ruke? Komar ili muxa?*
 what there at you on arm mosquito or fly
 ‘What is it on your arm? A mosquito or a fly?’ (Houtzagers 2003:203)
- (33) *Kto eto tebj ukusil?*
 who this you bit
 ‘[Looking at a swelling on someone’s hand clearly caused by an insect bite:] What stung you? (e.g., a wasp, a bee, etc.)’

I have also found an example of *kto* being used with grammatically inanimate nominals denoting biologically animate entities, as in (34), where the question is about a pot plant *anturium* ‘anthurium’.

Russian

- (34) [A: [...] *izvinite, stalo interesno: Vy odin anturium prodajte ili dva? Sprašivaju potomu, čto*
na 1-j fotografii anturium Andre, a na vtoroj – anturium
 on first photo anthurium PROP.GEN and on second anthurium
Šercera. Tak kogo vy prodajte? [B: *Prodavalsja anturium na*
 PROP.GEN so who.ACC you sell
[pervom foto, na vtorom prosto pokazan primernyj cvet cvetka. No v ljubom slučae anturium uže kupili.]
 ‘[(From a web forum, where people sell, buy and exchange pot plants) A:
 [...] sorry, I’m just interested: Do you sell one anthurium or two? I ask
 you because] on the first photo, there is *Anthurium andreanum* (lit.:
 ‘anthurium of André’) and on the second one *Anthurium scherzerianum*
 (lit.: ‘anthurium of Scherzer’). So, what (lit.: ‘who?’) are you selling? [B:
 It was the anthurium on the first photo that was on sale, the second photo
 was just to show the approximate colour of the flower. But anyway, the
 anthurium is already sold.]’⁶²

⁶¹ In this respect, recall the situation in Nganasan mentioned in Section II.4.1.3.

⁶² <http://hghltd.yandex.com/yandbtm?url=http%3A%2F%2Fwww.flowersweb.info%2Fforum%2Fread.php%3FFFID%3D10%26TID%3D54023%26MID%3D959013&text=%22%F2%E0%EA>

I suppose that it has been possible to use *kto* here because of the personal proper names in the names of the plants, viz. *anturium Andre* lit. ‘anthurium of André’ and *anturium Šercera* lit. ‘anthurium of Scherzer’.

Finally, in certain contexts, some grammatically animate nominals denoting biologically inanimate entities may be questioned with *kto* as well, as in (35) and (36).

Russian

- (35) [A:] *Eto Bol'shaja Medvedica, eto Lebed'...* [B:] *A eto kto?*
 this big she-bear this swan and this who
 ‘[Pointing to constellations on a sky map, A:] This is the Great Bear, this is Cygnus... [Pointing to another constellation, B:] And what is this?’
- (36) *Valeta b'jot dama, a damu kto? Korol'?*
 jack.ACC beats queen and queen.ACC who king
 ‘[About a card game:] A jack can be trumped by a queen, and what trumps a queen? A king?’

My impression is that the use of *kto* in (35) and (36) is possible only because the grammatically animate nominals involved also have as their original and primary meaning denotation of biologically animate things.⁶³ Therefore, I do not consider this and similar cases as a lack of differentiation between ‘who?’ and ‘what?’.

Note, however, that in other languages with grammatical animacy the situation may be somewhat different. Thus, a language with grammatical animacy may also allow for its animate non-selective interrogative pronominal to be used in questions about biologically inanimate things, when these are denoted by grammatically animate nominals which at the same time do not denote any

%20%EA%EE%E3%EE%20%E2%FB%20%EF%F0%EE%E4%E0%E5%F2%E5%22&reqtex
 t=(%22!%2B%F2%E0%EA%3A%3A636%20%20!%2B%EA%EE%E3%EE%3A%3A6336%2
 0%20!%2B%E2%FB%3A%3A540%20%20!%EF%F0%EE%E4%E0%E5%F2%E5%3A%3A1
 596091%22)%2F%2F6&dsn=0&d=2736300&sh=3&sg=37&isu=1 (created on 04.04.2007,
 retrieved on 18.05.2007).

⁶³ Thus, I suppose that *kto* may be used in (35) only if one assumes that constellations are normally called with grammatically animate nominals (which in fact, at least for the constellations of the northern hemisphere, would be a correct assumption, indeed). In (36), *kto* is preferable to *čto* because it is immediately followed by *korol'* ‘king’, which is grammatically an animate nominal denoting here an inanimate thing but which as its primary meaning, denotes a person. Note in this respect that *kozyr'* ‘trump’, which is often treated as grammatically animate, can be questioned only with *čto*, as in *Čto kozyr'?* ‘What’s trump?’. Apparently, the reason is that it does not have denotation of any biologically animate entity as its primary meaning.

biologically animate thing. Consider, for instance, (37) from Maliseet-Passamaquoddy where the animate interrogative pronominal *wen* is used with an animate nominal *lahtu* ‘trump’, which otherwise does not denote any biologically animate entity (cf. also Section III.7.2).⁶⁴ In origin, the word *lahtu* is a loan from French, viz. *l’atout* ‘the trump’ (cf. Francis & Leavitt 2007, under *lahtu*).

Maliseet-Passamaquoddy/Malecite-Passamaquoddy (Algonquian; Canada & USA)

- (37) *wen-ik lahtu-wok, tolepsis-ok kosona kalus-iyik?*
 IPW.AN-AN.PL trump-AN.PL club-AN.PL or diamond-AN.PL
 ‘What’s trump (lit.: ‘trumps’), clubs or diamonds?’ (Francis & Leavitt 2007, under *kalus*)

I prefer to treat such examples as a lack of differentiation between ‘who?’ and ‘what?’ because they cannot be accounted for by appealing to denotation of biologically animate things as the original and/or primary meaning of the nominals involved.

⁶⁴ Compare, in this respect, the Russian word *kozyr* ‘trump’ discussed in the preceding footnote.

5 Non-prototypical combinations of values: concluding remarks

In this chapter, I have examined in detail (i) the choice of ‘what?’ in KIND-questions (Section II.2), (ii) the choice of ‘who?’ in NAME-questions (Section II.3), and (iii) the choice of ‘who?’ in ANIMATE-questions (Section II.4). I have also discussed various constructional, semantic, areal and genetic preferences and restrictions that I found in the relevant languages as regards the choice between ‘what?’ and ‘who?’.

All in all, I have unequivocal data⁶⁵ for the three non-prototypical combinations of values for 192 languages (cf. Appendix D). This sample has been obtained through a non-controlled reduction of my global sample. That is, the reduction has been largely a matter of chance. As most of my data on the resolution strategies in cases of non-prototypical combinations of values comes from a questionnaire, the languages that ultimately made it into this reduced sample are mostly languages for which I found specialists that were willing *and* able to provide the relevant information on the three non-prototypical combinations of values. Although my sample is somewhat skewed as regards areal and genetic distribution, it seems to be diverse enough to be representative of the variation in the languages of the world. When compared to a variety sample of the same size constructed by Rijkhoff *et al.* (1993:186), the major positive deviations (in terms of percentage) in my sample occur for the Altaic, Uralic-Yukaghir and Indo-Hittite languages of Eurasia, while the major negative deviations occur in Indo-Pacific and Australian languages.⁶⁶ Based on the attested areal and genetic distribution of the languages recurring to the same resolution strategies of the three non-prototypical combinations of values in my global sample (which includes many gaps) (cf. Sections II.2.3, II.3.3-II.3.4, II.4.1.3), it may be supposed that if we eliminate the aforementioned deviations, this may involve, on the one hand, (i) an increase in the number of languages with ‘who?’-dominance in NAME-questions (i.e., [thing + identification (+ proper

⁶⁵ *Unequivocal* here means that I have answers in terms of *yes* or *no* rather than *possible*, *not clear* or *no information*. To be more precise, since preferences for a particular resolution (or avoidance) strategy in cases of non-prototypical combinations of values may vary according to contexts, *unequivocal* implies that in a given language, either (i) it is possible to use ‘what?’ in KIND-questions (or respectively, ‘who?’ in ANIMATE-questions or NAME-questions) in at least one context, or (ii) it is not possible to use ‘what?’ in KIND-questions (or respectively, ‘who?’ in ANIMATE-questions or NAME-questions) in any context.

⁶⁶ The 200-language sample in Rijkhoff *et al.* (1993:186) is constructed using Ruhlen’s (1987) classification of the languages of the world. For consistency sake, I also use it here, but it is not used elsewhere in the present study.

name)]), and on the other hand, (ii) a decrease in the number of languages with ‘what?’-dominance in KIND-questions (i.e., [person + classification (+ common noun)]). The number of languages with ‘who?’ in ANIMATE-questions (i.e., [animate thing + classification (+ common noun)]) will probably become somewhat lower in terms of percentage, since this kind of languages does not appear to be common in Australia and New Guinea.

Leaving aside the complications raised by ANIMATE-questions, the two major types of non-prototypical combinations of values, i.e. those in KIND-questions and in NAME-questions, and their resolution strategies, i.e. ‘who?’-dominance and ‘what?’-dominance, yield a four-way typology of languages, as shown in Table 1. Languages of type 1 show a general preference for the interrogative ‘who?’ when an interrogative has to be selected for a question that involves non-prototypical combinations of values on the parameters ENTITY TYPE and TYPE OF REFERENCE (& EXPECTED ANSWER), whereas languages of type 4 show a general preference for ‘what?’. In languages of type 2 the parameter ENTITY TYPE is decisive in the choice of an interrogative pronominal, whereas in type 3 the decisive parameter is TYPE OF REFERENCE (& EXPECTED ANSWER).

Table 1. The primary (four-way) typology of ‘who?’/‘what?’-dominance in cases of non-prototypical combinations of values

	KIND-questions [person + classification (+ common noun)]	NAME-questions [thing + identification (+ proper name)]	Prominence
1	‘who?’	‘who?’	‘who?’
2	‘who?’	‘what?’	ENTITY TYPE
3	‘what?’	‘who?’	TYPE OF REFERENCE
4	‘what?’	‘what?’	‘what?’

Adding the non-prototypical combination of values found in ANIMATE-questions in principle raises the number of logically possible types to eight. However, given that the definition of the entity types [person] and [thing] set forth in Section I.2.4 classifies non-human animates as [things] by default, I argue that the choice for ‘what?’ in ANIMATE-questions is the default choice. This default choice in ANIMATE-questions never calls for the definition of an extra type. Therefore, languages of type 1 and 2 come in two subtypes, depending on whether they assign non-human animates to the value [person] or not. Thus, only two types have to be added, viz. those in which ANIMATE-questions select ‘who?’, whereas KIND-questions select ‘what?’ (types 5 and 6). Table 2 presents the full typology.

Table 2. The full typology of ‘who?’/‘what?’-dominance in cases of non-prototypical combinations of values

		KIND-questions	NAME-questions	ANIMATE-questions	Prominence
1	a	‘who?’	‘who?’	‘who?’	‘who?’
	b			(‘what?’)	
2	a	‘who?’	‘what?’	‘who?’	ENTITY TYPE
	b			(‘what?’)	
3		‘what?’	‘who?’	(‘what?’)	TYPE OF REFERENCE
4		‘what?’	‘what?’	(‘what?’)	‘what?’
5		‘what?’	‘what?’	‘who?’	mixed (4/2a)
6		‘what?’	‘who?’	‘who?’	mixed (3/2a)

Table 3 represents the distribution of the six types (and two subtypes) among the languages of the reduced sample.

Table 3. The distribution of the languages of the reduced sample in terms of the full typology of ‘who?’/‘what?’-dominance in cases of non-prototypical combinations of values

Type	Number	% of total (192)	Languages (examples; cf. also Appendix D)	
1	a	5	3%	Barasana (Tucanoan), Swahili (Bantu), Trió (Carib)
	b	20	10%	Alawa (Maran), Candoshi (Isolate), Gweetaawu Dan (Eastern Mande)
2	a	1	1%	Russian (Slavic)
	b	29	15%	Budukh (Lezgitic), Daba (Chadic), Chickasaw (Muskogean)
3		42	22%	Eton (Bantu), Kaxarari (Panoan), Savosavo (East Papuan)
4		93	48%	Arabela (Zaparoan), Bisa (Eastern Mande), English (Germanic)
5		1	1%	Kathmandu Newar (Bodic)
6		1	1%	Avam Nganasan (Samoyedic)

Given that preferences for a particular resolution (or avoidance) strategy in the case of a non-prototypical combination of values may vary according to contexts, counts have been made from the perspective discussed in Section II.1. Thus, I have counted a given language as a language using ‘what?’ in KIND-questions if ‘what?’ is possible in at least one context (cf. Section II.2.1), otherwise the language has been counted as a language using ‘who?’ in KIND-questions.

Similarly, I have counted a given language as a language using ‘who?’ in NAME-questions if ‘who?’ is possible in at least one context (cf. Section II.3), otherwise the language has been counted as a language using ‘what?’ in NAME-questions. Finally, a given language was counted as using ‘who?’ in ANIMATE-questions if ‘who?’ is possible in at least one context (cf. Section II.4), otherwise the language has been counted as a language using ‘what?’ in ANIMATE-questions. The percentages in the third column (% of total) have been rounded to the closest whole number.

Example (1) illustrates the only language of type 6 in my sample, viz. Avam Nganasan. (1a) is an example of a KIND-question formulated with ‘what?’, (1b) is an example of a NAME-question formulated with ‘who?’, and (1c) is an example of an ANIMATE-question where ‘who?’ is used.

Avam Nganasan (Uralic, Samoyedic; Russia; Valentin Goussev, p.c.)

- (1) a. [A:] *takəə-tə* *bəjkaʔa* *maaʔ* [B:] *setəgəə*
 DEM-2SG.GEN husband what director
 ‘[A:] What is her husband? [B:] A director’
- b. *ʃili* *nʔim-tiʔ*
 who name-3SG
 ‘What is his name? (lit.: ‘Who is his name?’)’
- c. *ʃili* *tənə* *sakələ-hu* *dʔüδü-mtəʔ*
 who 2SG bite-Q.PST arm-ACC.2SG
 ‘[Someone walks into a room with a swelling on his arm, a wound. A:]
 What bit you in your arm?’

The English translations of the examples in (1) illustrate that English is a language of type 4, since the translation of (1a) is an example of a KIND-question formulated with ‘what?’, the translation of (1b) is an example of a NAME-question formulated with ‘what?’, and the translation of (1c) is an example of an ANIMATE-question involving ‘what?’.

The comparative frequency of the types in Table 3 in terms of percentage is summarized in (2). (2a) takes into account only the four basic types, viz. types 1-4 (cf. Table 1 above) without distinguishing between the subtypes of types 1 and 2. (2b) adds types 5 and 6 to (2a). Finally, (3) takes into account all types and subtypes. In brackets, I give the respective percentages.⁶⁷ I have arbitrarily chosen a $\geq 10\%$ difference in frequency as the threshold, which I mark by means of the symbol ‘>’. Less significant differences in frequency, viz. in the interval between 10% and 5%, are marked with the symbol ‘>’. Even smaller differences or lack thereof are marked by a comma.

⁶⁷ The sums of the percentages are not equal to 100% because the percentages have been rounded.

- (Y2) a. type 4 (48%) > type 3 (23%) › type 2 (16%), type 1 (13%)
 b. type 4 (48%) > type 3 (22%) › type 2 (16%), type 1 (13%) > type 5 (1%), type 6 (1%)
 c. type 4 (48%) > type 3 (22%) › type 2b (15%) › type 1b (10%) › type 1a (3%), type 2a (1%), type 5 (1%), type 6 (1%)

Alternatively, we may count the differences between the percentages for each type (and subtype) and the average for each of the three counts, as summarized in (3). Thus, the average for (2a) is 25% (i.e., 100% divided by 4 types), which means that when we count only the four basic types, type 4 (48%) is 23% above the average, type 3 (23%) is 2% below the average, type 2 (16%) is 9% below the average, and type 1 (13%) is 12% below the average, as summarized in (3b). The average for (2b) is 17% (100% divided by 6 types), the relation between the percentages in (2b) and the average for (2b) is summarized in (3b). The average for (2c) is 13% (100% divided by 8 (sub)types), the relation between the percentages in (2c) and the average for (2c) is summarized in (3c). The (sub)types with the frequency above the average are highlighted in bold.

- (3) a. **type 4** (+23%) > **AVERAGE** > type 3 (-2%) › type 2 (-9%), type 1 (-12%)
 b. **type 4** (+31%) > **type 3** (+5%) › **AVERAGE**, type 2 (-1%), type 1 (-4%) > type 5 (-16%), type 6 (-16%)
 c. **type 4** (+35%) > **type 3** (+9%) › **type 2b** (+2%) › **AVERAGE**, type 1b (-3%) › type 1a (-10%), type 2a (-12%), type 5 (-12%), type 6 (-12%)

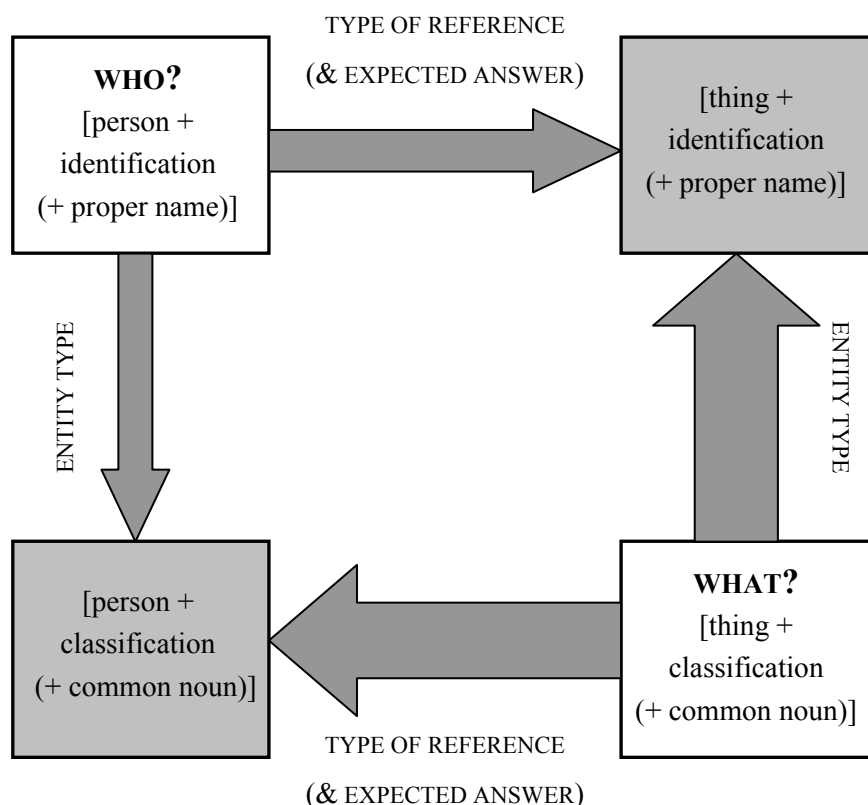
Let us discuss the results summarized in (2) and (3) in more detail. To begin with, note that type 2b, which, judging from most grammatical descriptions, could have been assumed to be the default case, is by no means the highest in frequency (2c) and it is only slightly higher in frequency than the average (3c).

Second, note the overall minimal frequency of the types involving ANIMATE-questions formulated with ‘who?’, viz. subtype 1a, subtype 2a, type 5, and type 6, cf. (2b-c) and (3b-c). Possible reasons for the rarity of ANIMATE-questions using ‘who?’ have been discussed in Section II.4.1.3.

Third, what strikes the eye is the very high frequency of type 4, which involves ‘what?’-prominence. In addition, we may also note that the type with the second highest frequency is type 3 involving TYPE OF REFERENCE-prominence. What the two types, 4 and 3, share is ‘what?’-dominance in KIND-questions. In my reduced sample, ‘what?’-dominance in KIND-questions is found in around 70% of the languages, as opposed to around 30% for ‘who?’-dominance. In terms of frequency, type 3 is followed by type 2 and type 2 shares with type 4 ‘what?’-dominance in NAME-questions. In my reduced sample, ‘what?’-dominance in NAME-questions is found in around 65% of the languages,

as opposed to around 35% for ‘who?’-dominance. These asymmetries between ‘who?’- and ‘what?’-dominance is schematized in Figure 1 (based on Figure II.1:5).

Figure 1. Asymmetries between ‘who?’- and ‘what?’-dominance in the reduced sample



It is not implausible that the asymmetries between ‘who?’- and ‘what?’-dominance represent some universal preference in human languages. This preference could be explained in terms of semantic markedness, for instance. That is, semantically, ‘who?’ may be considered as more marked than ‘what?’. Apparently, the main reason for this would be the more marked status of the category PERSON as compared to the category THING (cf. Section I.2.4; as well as Section I.4.2.3.3.1 on “no conjecture”-contexts).

Although on the whole this explanation of the asymmetries between ‘who?’- and ‘what?’-dominances in the reduced sample may be adequate, it may also be hypothesized that at least to some extent the high degree of asymmetry observed is enhanced by the dominant position of type 4 languages such as English, French, Spanish and Portuguese in many regions of the world. The latter hypothesis may be supported by the following facts. To begin with, interrogative pronominals are lexical elements and lexical meanings are easily borrowed. That

is, in many cases, the preference for ‘what?’-dominance may have been taken over from a dominant Western European language.⁶⁸ It may also be interesting to note that older Indo-European languages such as Latin and Ancient Greek seem to have been of type 2(b) (this issue requires further investigation). Furthermore, at least as far as the preference for ‘what?’-dominance in KIND-questions is concerned, recall that as discussed in Section I.4.2.3.3.4, in adult-to-infant talk, the same Western European languages often appear to prefer ‘who?’ to ‘what?’. That is, in adult-to-infant talk these languages belong to type 2, which seems to suggest that there may be a shift from type 2 to type 4 at a certain phase in the process of language acquisition. Finally, when interpreting the numerical data of the reduced sample, we should not forget that as discussed in the beginning of the present section, my reduced sample is somewhat skewed in favour of the Eurasian languages to the disadvantage of the languages of Australia and New Guinea. A possible correction of this areal bias may make the transitions between the frequencies of different types less pronounced in some cases.

Besides counting the absolute frequencies of the different types (2) and their relation to the average (3), we may also compare the observed frequencies of the four major types (within the four-way typology as presented in Table 1) with the frequencies of the same four types statistically expected on the assumption of fixed margins, i.e. on the assumption that the totals for the rows and the columns in a tetrachoric table such as Table 4 do not change. Table 4 summarizes the attested number of occurrences of types 1-4 in the reduced sample (as provided in Table 3). Numbers ❶, ❷, ❸, and ❹ indicate the type.

Table 4. The attested number of occurrences of types 1-4 in the reduced sample

		KIND-questions		rows total
		‘who?’	‘what?’	
NAME-questions	‘who?’	25 ❶	43 ❸	68
	‘what?’	30 ❷	94 ❹	124
columns total		55	137	192

⁶⁸ Even within Europe we can find clear examples of the influence of the dominant language on a minority language as regards the preference for ‘who?’- or ‘what?’-dominance. For instance, Upper Sorbian, a Slavic language spoken in the east of Germany, belongs to the same type 4 as German (Leńska Scholze, p.c.), whereas the closely related Polish (as well as most other Slavic languages) belongs to type 2.

Table 5 provides (in bold) the expected values (by chance alone) for the data from Table 4. In brackets, the deviations from the actual values are given. The expected value for cell X in column C and row R can be calculated with the following formula: $\text{Expected}_X = (\text{Total}_C \times \text{Total}_R) / \text{Total}_{\text{grand}}$ (cf. Janssen *et al.* 2006:424). In the case of type 1, for instance, this results in $(68 \times 55) / 192 = 19.5$ (rounded).

Table 5. The statistically expected values for the data from Table 4

		KIND-questions		
		'who?'	'what?'	rows total
NAME-questions	'who?'	19.5 (+5.5) ①	48.5 (-5.5) ③	68
	'what?'	35.5 (-5.5) ②	88.5 (+5.5) ④	124
columns total		55	137	192

Table 5 shows that type 1 and type 4 are more frequent than could be expected due to pure chance, whereas type 2 and type 3 are less frequent. This may account for the fact that there is such a big gap in frequency between type 4 and type 3, on the one hand, and only a small gap between type 1 and type 2, on the other hand.⁶⁹

The areal distribution of the languages involved has already been discussed in the respective sections on KIND-questions in Section II.2, NAME-questions in Section II.3, and ANIMATE-questions in Section II.4. Finally, given that my reduced sample is rather small, there seems to be little sense in counting the types in terms of genera and macro-areas.

By way of conclusion, a few words should be said on the issue of the diachronic stability of resolution strategies (and their combinations) in cases of non-prototypical combinations of values. Given that for most languages no

⁶⁹ Different statistical tests can be applied to the data in Table 4 to determine whether the deviations from the expected values, as presented in Table 5, are significant. Significant here would mean that there is a true interaction between the preferences for 'who?'-dominance or 'what?'-dominance in KIND-questions and NAME-questions, so that (i) 'who?'-dominance in KIND-questions really tends to go together with 'who?'-dominance in NAME-questions, and (ii) 'what?'-dominance in KIND-questions really tends to go together with 'what?'-dominance NAME-questions. However, my limited knowledge of statistics does not allow me to make an educated choice between the different tests (for a discussion see, e.g., Cysouw 2003; Maslova 2003; Dryer 2003; Janssen *et al.* 2006).

historical data on resolution strategies in cases of non-prototypical combinations of values are available, our judgements can only be based on comparison of related languages. In my sample, there is often substantial variation between closely related languages, which implies that diachronically, preference for a given resolution strategy (and their combinations) in cases of non-prototypical combinations of values is not a particularly stable phenomenon.

III Lack of differentiation

1 Introduction

In what follows, I will discuss various languages that appear to allow for a lack of differentiation between ‘who?’ and ‘what?’. The languages are grouped geographically in seven areas: (i) Africa and the Middle East (Section III.2), (ii) Eurasia (Section III.3), (iii) Southeast Asia and Oceania (Section III.4), (iv) New Guinea (Section III.5), (v) Australia (Section III.6), (vi) North America (down to Panama in the south and excluding the islands of the Caribbean; Section III.7), (vii) South America (Section III.8). Within these areas the languages are organized genetically. For each language I will first try to determine whether we can truly speak about a lack of differentiation between ‘who?’ and ‘what?’, and if so, what the origins (if any) of the lack of differentiation between ‘who?’ and ‘what?’ could be in each particular case. However, what will matter for me most here is whether the same interrogative can be used as both ‘who?’ and ‘what?’ and not whether and how in a given language, it is possible to disambiguate between the two meanings by some other means.

Given the wide scope of the present study, it has proved impossible to consider every area, every linguistic group and every particular language in the same degree of detail. In general, the order of areas in which the languages under discussion are organized reflects the order in which I have investigated them. This implies that the areas closer to the end of the list, such as North and South America, have been investigated later and in less detail than the areas in the beginning of the list, which were investigated first. The only major discrepancy between the two orders is represented by Indo-European languages, which are considered in Section III.3 on Eurasia but were investigated as the last ones. Such a “belated” interest in Indo-European is due to the following reasons. Given that (i) on the whole, Indo-European languages are among the best studied in the

world, and that (ii) there are only two Indo-European languages commonly mentioned in the typological literature to allow for a lack of differentiation between ‘who?’ and ‘what?’, viz. Latvian and Lithuanian, the situation with Indo-European languages seemed to be clear enough to deserve a more detailed study so that their analysis was left for the end. Although my expectations proved to be misplaced, the limitations of time have not allowed me to examine Indo-European languages in the same degree of detail as many of the other Old World languages.

Availability *and* usability of data have also been an important factor as regards the degree of detail to which a given language, linguistic group or area have been investigated. Thus, some more philologically oriented traditions are not easily accessible to the “non-initiated”, so that getting acquainted with them has required more time than strictly speaking, was available. Furthermore, if one wants to establish the origins of a lack of differentiation between ‘who?’ and ‘what?’ in a given language in the most plausible way, it may often be desirable to take all other related languages into consideration as well. However, in many cases it has proved to be rather difficult to put this requirement in practice.

2 Africa and the Middle East

In this section, I will discuss the languages presumably allowing a lack of differentiation between ‘who?’ and ‘what?’ that are spoken in Africa and the Middle East. The section is organized as follows. In Section III.2.1, I will examine Niger-Congo languages, in Section III.2.2 “Khoisan” languages and in Section III.2.3 Afro-Asiatic languages.

2.1 Niger-Congo languages

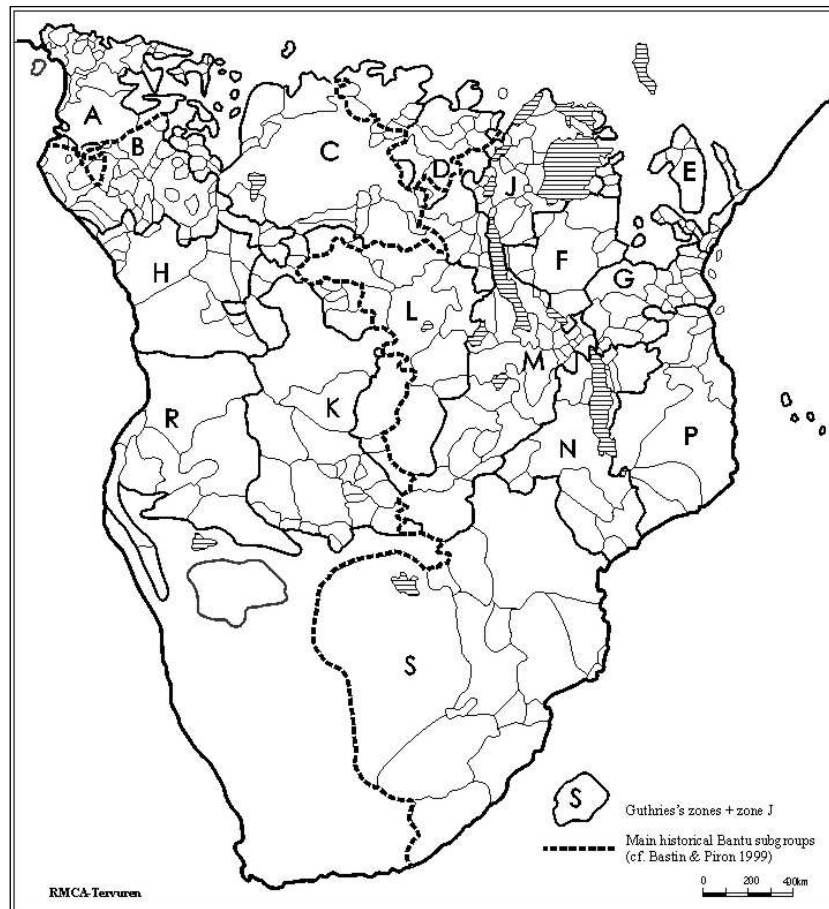
The Niger-Congo phylum is the largest linguistic group in Africa. Although the lower-level subgroups of Niger-Congo are clear, the exact relations between them remain a matter of discussion. The languages allowing a lack of differentiation between ‘who?’ and ‘what?’ have been found in two subgroups of Niger-Congo, viz. among the Bantu languages of Central Africa (Section III.2.1.1) and the Atlantic languages of West Africa (Section III.2.1.2). The lack of differentiation between ‘who?’ and ‘what?’ in the two groups has clearly evolved independently in the two groups (and due to different reasons).

2.1.1 *Bantu languages*

2.1.1.1 Introduction

Bantu languages form a low-level subgroup within the Benue-Congo branch of the Niger-Congo phylum. There are some 500-600 Bantu languages, which are spoken throughout Central, Eastern and Southern Africa. Traditionally, for reference purposes all Bantu languages are subdivided into 15-16 zones, named from A to S (with some gaps). A given Bantu language is usually referred to by a combination of an upper-case letter for the zone and a two digit number for the group and the language (further extensions are possible). Group numbers are multiples of ten. For instance, Swahili can be referred to as G42 or, with reference to the group only, as G40. Currently, there are at least three different reference systems in use (cf., e.g., Maho 2001, 2002 for an overview). Consequently, the same language may be referred to with somewhat different codes in different sources. For convenience sake, I use the codes found in the Ethnologue (Gordon 2005). Note that the Ethnologue seems to provide group numbers only.

It should be emphasized that strictly speaking, Bantu zones should be viewed as areal rather than genetic groupings. The distribution of the Bantu zones is presented on Map 1.

Map 1. Bantu zones (according to the Tervuren system)¹

Several Bantu languages of zone C have been found to allow for a lack of differentiation between ‘who?’ and ‘what?’. As far as I can judge from the data used in Bastin *et al.* (1999) and my additional data, which together probably cover more than 90% of the Bantu languages, zone C must be the only zone with such languages.² Depending on the way one counts languages, the number of the zone C languages allowing for a lack of differentiation between ‘who?’ and ‘what?’ may vary from 3 to at least 10, and perhaps even much more. In my database I have (arbitrarily) included 5 idioms: (i) *ndè/nê* ‘who?, what?’ Mbosi/Mboshi (some variants; Bantu C30; Congo; Bastin *et al.*’s 1999 database), (ii) *nâ* ‘who?, what?’ Tetela (some variants; Bantu C80; Democratic Republic of the Congo; Bastin *et al.*’s 1999 database), (iii) *ná* ‘who?, what?’

¹ © The Royal Museum of Central Africa in Tervuren (Belgium).

² Bastin *et al.*’s (1999) database contains two languages outside of zone C marked as using a single interrogative ‘who?, what?’, viz. Soli (Bantu M60; Zambia) and a variety of Nyanja (Bantu N30; Malawi). However, this characterization has proved to be a misprint for Nyanja (cf. Missionários da Companhia de Jesus 1964) and is probably due to a misinterpretation in the case of Soli (‘who?’ and ‘what?’ seem to be distinguished through gender prefixes).

Mongo(-Nkundo) (Bantu C70; Democratic Republic of the Congo; Hulstaert 1938, 1961, 1965, 1966, 1993), (iv) *ɛ́nɔ* Ntomba-Inongo (Bantu C70; Democratic Republic of the Congo; Bastin *et al.*'s 1999 database), (v) *ɛ́ɔ* Bolia (Bantu C40; Democratic Republic of the Congo; Mamet 1960:35). The last three idioms are sometimes considered as “dialects” of one language, Mongo. The term *Mongo language* is a somewhat artificial construct created by the Flemish missionary Gustaaf Hulstaert to refer to a large dialect cluster, or rather, a set of closely related languages, spoken in the vast region between the Kasai and the Congo Rivers.³ The Mongo variety that served as *dialecte de base* for Hulstaert and that his publications primarily describe is Nkundo. Nkundo is spoken in the Democratic Republic of the Congo in the region around Boteka where the Flandria Mission of Hulstaert was situated. Given that Mongo is by far the best described language of zone C, the following discussion will be confined to Mongo.

2.1.1.2 Mongo (Bantu C70)

Mongo, as described by Hulstaert (1938, 1961, 1965, 1966, 1993), has two interrogative pronominals, *ná* ‘who?, what?’ and *é* ‘what?’, as in (1-2).

Mongo (Hulstaert 1965:533)⁴

- (1) a. *á-kela ná?*
AG1-do IPW
‘What is he doing?’
b. *á-kela é?*
AG1-do IPW
‘What is he doing?’
- (2) a. *ilɔmbɛ yǎ ná?*
G19.house AG19.CON IPW
‘a house of whom?, whose house?’ or ‘a house of what?, a house for

³ The “creation” of the Mongo language is particularly interesting from historic and sociolinguistic points of view. Hulstaert can by right be called the first Mongo “nationalist”. He aspired to create one big language on the basis of the Bantu idioms spoken in the region between the Kasai and the Congo Rivers hoping that such a big language will be able to rival French as a national language of the Belgian Congo (see Van de Velde 1999).

⁴ In Mongo examples, elisions are not represented and only genders and agreement patterns are regularly marked. In examples from Bantu languages in the present section, verbal categories are normally not glossed, since the verbal morphology of Bantu languages is rather complicated both formally and structurally but bears no relevance for the present discussion. Moreover, glosses are usually lacking in the sources. Numbers of agreement patterns and genders in all examples from Bantu languages in this section follow the traditional Bantuist conventions, as found in Meeussen (1967).

what (purpose, thing)?’

- b. *ilɔmbɛ* *yǎ* *é?*
 G19.house AG19.CON IPW

‘a house of what?, a house for what (purpose, thing)?’

Other interrogative proforms are *nkó* ‘where?’, *mó* ‘how?’, *ngámó* ‘how?; what quality of [N]?, what kind of [N]? (asks for a description)’,⁵ AG-*ngá* ‘how many of?’, AG-*lénkó* ‘which one?’ (< AG-COP where?). The interrogatives *ná* and *é* are also encountered with the following secondary meanings: *ná* can be used as an exclamative postnominal modifier, something like ‘what a [day, view, person, etc.]!’, and *é* can sometimes be used instead of *nkó* ‘where?’ and as a polar question marker.

In what follows, I will first discuss nominal uses of *ná* and *é* in Section III.2.1.1.2.1. Secondary uses will be considered in Section III.2.1.1.2.2. In Section III.2.1.1.2.3, I will argue that probably, the lack of differentiation between ‘who?’ and ‘what?’ in Mongo is due to the fact that the interrogative *ná* ‘who?, what?’ goes back to a selective interrogative pronominal meaning ‘which one? (person or thing)’, and ultimately the locative interrogative ‘where?’.

2.1.1.2.1 The interrogatives *ná* and *é*: nominal uses

Syntactically, *ná* ‘who?, what?’ and *é* ‘what?’ generally behave like nouns. They can be objects, as in (1) above, also prepositional ones, as in (3).

Mongo (Hulstaert 1938:80)

- (3) a. *lá* *ná?*
 with IPW
 ‘with whom?’ or ‘what for?, why?’
 b. *lá* *é?*
 with IPW
 ‘what for?, why?’

Judging from the texts I analyzed, *é* ‘what?’ is very rarely used. One of the few contexts where it is much more frequent than *ná* is in combination with the preposition *lá* ‘with’, as in (3), to ask ‘what for?, why?’. The words *ná* ‘who?, what?’ and *é* ‘what?’ can also function as a modifying element in a connective construction, as in (2). Moreover, they can be used as the second element of a nominal modifying construction without a connective,⁶ as in (4), which can be

⁵ The interrogative *ngámó* has the meaning ‘what quality of?, what kind of?’ when used as the modifying element of a connective construction similar to that in (2).

⁶ The term *connective* is used in Bantu studies to refer to a functional word used to join two nominals in a construction where one nominal modifies the other in some way.

compared to (5).

Mongo

- (4) *(lá) ntsína ná/é?*
 with G9.reason IPW/IPW
 ‘why?’ (Hulstaert 1938:80)
- (5) a. *baíso jǎle*
 G2.eyes G5.ferocity
 ‘glaring eyes’ (Hulstaert 1966:106)
- b. *eyenga bonkúnju*
 G7.week G5.totality
 ‘the whole week’ (Hulstaert 1966:102)

Neither *ná* nor *é* can function as (i) subjects, (ii) heads of connective constructions, as they might have been in phrases like *what of John (have you seen yesterday)?* (e.g., John’s house) or *who of the villagers (can afford it)?*, or (iii) head one of the few Mongo adjectives.⁷ This implies that although *ná* and *é* are similar to nouns in other respects, they cannot occur in any position where they control agreement, which means that it is impossible to determine their gender and they should probably be best viewed as genderless.

In Mongo, which has SVO order, interrogatives are minimally clause-final, and most typically sentence-final (Hulstaert 1966:876-877). When *ná* ‘who?, what?’ and *é* ‘what?’ question the subject of a declarative clause, the following strategy has to be used. The predicate of a corresponding declarative clause is realized as a headless relative clause, while the interrogative is postposed to it into sentence-final position, as in (6).

Mongo (Hulstaert 1965:144)

- (6) *ǒ-kelaki ná?*
 AG1.REL-did IPW
 ‘Who has done it? (lit.: ‘The one who has done [it] (is) who?’)’

The relative can be analyzed as the subject of the question and the interrogative as the predicate (cf. Hulstaert 1966:470). Thus, a question about subject like *Who did it?* is literally something like *The one who has done it is who?*. Similarly, one cannot say *What/which [N] does this?* using a connective-less nominal modifying construction with the interrogative as a modifier. Instead, a paraphrase with a relative form has to be used, viz. *The [N] who does this is what/who/which one?*. The interrogative construction based on a relative clause is also available for

⁷ Hulstaert does not mention explicitly the possibility of using *ná* and *é* in the last two contexts, but the native speakers consulted clearly reject it. I am grateful to Honoré Vinck for checking this for me with native speakers.

questioning objects, but it appears to be used less frequently than for questioning subjects.

Only in certain “dialects” of Mongo, such as Bolia, do we sometimes find the interrogative pronominals questioning subjects in sentence-initial position (7).

Bolia (Bantu C40; Democratic Republic of the Congo; Mamet 1960:35, 125)

- (7) *ńɔ* *ǝ-tááká* *itóíbá* *bentómba*
 IPW AG1.REL-will.go to.go.to.steal G4.giant.pouched.rat
bě-lé *bɔ* *ńɔ?*
 AG4.REL-eat PL IPW

‘Who would ever go to steal the giant pouched rats, who would eat them?’

Note, however, that the verb in (7) is still in the relative form and the interrogative is repeated sentence-finally, but now in plural (regarding the plural form, see also the comment before example (10) below).

The use of the relative form in questions about subjects, as well as the sentence-final occurrence of interrogatives, results from the general principles of coding of information structure in Mongo: a topical element normally occurs before the predicate, while the focused element occurs after the predicate or, if it is part of the predicate, at the end of the predicate phrase (cf. Hulstaert 1966:424-482).⁸ Relativization can be conceived of as a means of topicalizing the predicate, and indirectly, focalizing the subject.⁹ It is also possible in Mongo to vary the position of (and according to Hulstaert 1966:470, the syntactic relations between) the relative form and its notional subject in order to express different kinds of focus. Whereas [notional subject + relative form] order usually implies

⁸ There is currently a project “Word order variation and wh-questions in Bantu” (<http://www.lucl.leidenuniv.nl/index.php3?m=&c=78>) at the University of Leiden dedicated specifically to the study of post-verbal placement of interrogative proforms in other Bantu languages. The project is situated within the generative framework and hypotheses proposed vary from “rightward movement” to “front[ing] of wh-phrases with subsequent remnant movement of the rest of the sentence” even further to the left.

⁹ According to Raymond Boyd (p.c.), focalization of subjects that are questioned is widespread in the Niger-Congo languages of Central Africa. For instance, in Zande (Niger-Congo, Ubangi; Central African Republic & Democratic Republic of the Congo) interrogatives may appear *in situ* or they may be focalized. However, when the interrogative pronominals ‘who?’ or ‘what?’ question the subject, they *must* be focalized. In fact, quite a few Bantu languages outside of Central Africa behave similarly as well. For instance, in Northern Sotho (Bantu S30; South Africa) and Venda (Bantu S20; South Africa) “interrogatives [...] may not function as straightforward *subjects* of verbs... [one] gets around this problem by using a copulative construction with a following relative” (Poulos & Louwrens 1994:376-377; Poulos 1990:451-452). See also Bearth (2003).

contrastive focus, as in (8a),¹⁰ the reverse order [relative form + notional subject] appears to be used for completive focus (Hulstaert 1966:472-473; see Dik 1989:282 for the notions of contrastive and completive focuses), as in (8b). The latter order is also used in answers on questions about subjects (Hulstaert 1966:471).

Mongo (Hulstaert 1966:472)

- (8) a. *ô nyama ô-wéi*
 just G9.animal AG1.REL-died
 ‘It is just an ANIMAL that died (and not a man).’
 b. *ô-wéi ô nyama*
 AG1.REL-died just G9.animal
 ‘The one who died is just an animal.’

Another interesting feature of questions about subjects, which they share with focus constructions as those in (8), is the agreement pattern of the relative form. Usually, it is AG1, irrespectively of whether the intended meaning is ‘who?’ or ‘what?’. In this respect, compare (6) to (9).

Mongo

- (9) [*Bisénga, betúbyaíso! Wolo nkína témpelo žyéngwéyá wolo,*]
 ô-lekí ná?
 AG1.REL-be.superior IPW
 ‘[Ye fools and blind:] for whether is greater, [the gold, or the altar that sanctifieth the gold]?’ (Matthew 23:17 in Hulstaert 1987)¹¹

When the presupposed referent is human and plural, agreement pattern AG2, the plural counterpart of AG1, is also possible. The interrogative *ná* is then typically

¹⁰ The pre-verbal occurrence of the interrogative in the Bolia example (7) is probably due to the same reason. Note that this question is rather rhetorical in nature. It is uttered by the slaves who are being tortured by their master who accuses them of having stolen five smoked rats. The slaves say that the only reason they are accused of theft is that they are slaves and that they would never go to their master’s place. Thus, they rhetorically ask their master using contrastive focalization: *Who (of the slaves) would ever (dare to) go to steal the rats?*, implying that no slave would ever do so. The second part of the question *bělé bó ná* is best interpreted as a separate question *Who (plural) would eat the rats?*. Since the smoked rats did disappear and somebody must have eaten them, the question is not rhetorical anymore and the interrogative occupies the regular final position.

¹¹ It may be argued that *ná* in (9) is selective, viz. ‘which one?’. Although in principle, this interpretation is possible, it is preferable, I believe, to interpret *ná* in (9) as ‘what?’, given that Mongo has a separate word for ‘which one?’, viz. *-lénkó*. In this respect, see also the discussion in Section I.2.3.

preceded by a plural proclitic *baa*,¹² as in (10).

Mongo (Hulstaert 1965:144)

- (10) *bǎ-kelákí baa ná?*
 AG2.REL-do PL IPW
 ‘Who has done it? (lit.: ‘The ones who did it are who?’)’

Hulstaert (1966:7) refers to the use of agreement pattern AG1 in (6) and (9) as “absence of agreement”, which is governed by a rather complex set of rules. With that Hulstaert means the use of agreement pattern AG1 even when the controller is not of gender 1 or no controller is present. Gender 1 is traditionally described as a “human” gender in descriptions of Bantu languages and since gender is determined through agreement, AG1 is also normally labelled as “human”. Still, in many Bantu languages this “human” agreement has been reported to be used with non-human and even non-animate controllers, such as for instance interjections or infinitives. Corbett calls this a perplexing choice in his typological study of gender (1991:208). However, Van de Velde (2006) points out that agreement patterns can have various functions, agreement in gender being only one of them. Thus, he convincingly argues that agreement pattern AG1 in Bantu can mark agreement not only with controllers that belong to gender 1 or have human (or animate) reference, but also those that “are not in need of referential disambiguation”, either because they are already inherently definite or because they cannot be disambiguated. Thus, in Mongo agreement pattern AG1 is used for agreement with generic referents (11b), agreement with indefinite referents (12b) and so-called “enforced agreement” (13).¹³ Agreement pattern AG1 in sentences like (6), (8b), and (9) appears to represent yet another similar usage: the marking of suspended reference (cf. Section I.2.1).

Mongo

- (11) a. *mpulú é-tónɡa júmbu*
 G9.bird AG9-build nest
 ‘The bird is building a nest’ (Hulstaert 1966:17)

¹² Hulstaert describes it as gender 2a prefix, but almost always writes it separately. A few other morphosyntactic peculiarities of *baa* seem to indicate that it is a plural word (in the sense of Dryer 1989), rather than a prefix. This might be a widespread phenomenon in Bantu (cf. Van de Velde forthcoming).

¹³ The term *enforced agreement* refers to the use of agreement markers when an atypical controller such as an interjection is used or no controller is possible at all. In a gender language, if an agreement target *can* agree then typically it *must* agree (Corbett 1991:204).

- b. *mpulú á-tóngá jumbu*
 G9.bird AG1-build nest
 ‘Birds build nests (general truth)’ (Hulstaert 1966:17)
- (12) a. *mbwǎ ě-owá*
 G9.dog AG9-died
 ‘The dog has died’ (Hulstaert 1966:18)
- b. *mbwǎ ě-owá*
 G9.dog AG1-died
 ‘A dog died’ (Hulstaert 1966:18)
- (13) *“nkína” á-foténe ekoli*
 maybe AG1-NEG.cut G7.discussion
 ‘“Maybe” does not end a discussion’ (Hulstaert 1966:22)

In some cases when there is no overt controller, it may still be possible to think of a potential controller. Consequently, the speaker may decide to use the agreement pattern corresponding to the potential controller rather than resort to the agreement pattern AG1. In such cases, one can probably speak about “exophoric” agreement (compare Corbett 1991:243-244). The potential controller will be normally thought of in basic-level terms (see Rosch 1978 on the latter notion). Thus, in Mongo, agreement pattern AG4 can be used when one thinks about the gender 4 words *besakó* ‘news’ and *belema* ‘works; problems’ (Hulstaert 1965:648), as in (14).

Mongo (Hulstaert 1965:648)

- (14) a. *bě-le ndá bolá ná?*
 AG4.REL-be in G3.village IPW
 ‘What is the news from the village?’
- b. *ă-owofola bé-umá bě-lekákí ndá bonanga*
 AG1-reminded.him AG4-all AG4.REL-passed in G3.tribe
 ‘He reminded him about everything that occurred in the tribe’

Agreement pattern AG5 can be used when one thinks about gender 5 word *jói* ‘action, event; matter; words’ (Hulstaert 1965:649), as in (15).

Mongo

- (15) a. *lĩ-kwě-i ndá bolá ná?*
 AG5.REL-happened in G3.village IPW
 ‘What happened in the village?’ (Hulstaert 1965:649)
- b. *ǰ-ólénaka jói lĩ-kwêkí ǎnko?*
 2SG-saw G5.event AG5.REL-happened there
 ‘Have you seen what happened there?’ (Hulstaert 1957:897)

However, in general, when there is no overt controller, agreement pattern AG1, or its plural counterpart, AG2 (i.e., enforced agreement) is used in Mongo much more frequently than other agreement patterns (Hulstaert 1965:644).

2.1.1.2.2 The interrogatives *ná* and *é*: other uses

In addition to the nominal uses discussed in Section III.2.1.1.2.1, the interrogatives *ná* and *é* can also have some non-nominal uses. Thus, *é* can function as a polar question marker, as in (16).

Mongo (Hulstaert 1965:533)

- (16) *ǎ-olúndola é?*
 AG1-came.back PQ
 ‘Is he back?’

In fact, this is its most frequent use (Hulstaert 1965:532). Development from ‘what?’ to a polar question marker is not particularly surprising. Very occasionally and probably only with verbs with appropriate semantics (movement, posture, etc., with a valency for a locative argument), *é* can be used instead of *nkó* ‘where?’, as in (17).

Mongo (Hulstaert 1965:533)

- (17) *ǎ-otswá é?*
 AG1-went where
 ‘Where has he gone?’

The interrogative *ná* can be used as an exclamative postnominal modifier, as in (18).

Mongo (Hulstaert 1966:167)

- (18) *lifokú ná!*
 G5.beauty IPW
 ‘What beauty!’

This usage can be considered to be an instance of a nominal modifying construction without a connective, as in example (5). That is, the development is from an interrogative *what (kind of) beauty?* to the exclamative *what (kind of) beauty!*, which is especially easy to conceive if one knows that exclamative intonation in Mongo is the same as interrogative intonation. Hulstaert even notes that with *ná* sometimes only the context helps to tell whether the sentence is interrogative or exclamative (1961:150).

The use of the same form both pronominally as ‘who?’ and attributively as ‘what [N]?, what kind of [N]?’ is attested elsewhere in Bantu, also in languages where ‘who?’ and ‘what?’ are otherwise different. Consider, for instance, the

interrogative *mang* from Tswana in (19), which otherwise is used as ‘who?’.

Tswana (Bantu S30; Botswana)

- (19) a. *thipa mang?*
 G9.knife IPW
 ‘what (kind of) knife?’ (Andy Chebanne, p.c.)
- b. *thipa e-fe?*
 G9.knife AG9-IPW
 ‘which knife?’ (Andy Chebanne, p.c.)
- c. *ke nako mang?*
 COP G9.time IPW
 ‘What time is it?’ (http://en.wikipedia.org/wiki/Tswana_language)

As in many other Bantu languages, the interrogative ‘who?’ in Tswana can also be used in NAME-questions (cf. Section II.3), as in (20).

Tswana (Bantu S30; Botswana; Andy Chebanne, p.c.)

- (20) b. *lefelo je le bidiwa mang?*
 G5.place AG5.this AG5 is.called who
 ‘What is this place called? (lit.: ‘Who is this place called?’)’ (e.g., Gaborone)
- b. *leina ja setlhare se ke mang?*
 G5.name AG5.CON G7.tree AG7.this COP who
 ‘What is the name of this tree? (lit.: ‘The name of this tree is who?’)’ (e.g., oak)

In Mongo the situation is similar (at least for personal proper names), as can be seen in (21), where *ná* cannot be replaced with *é* ‘what?’ and therefore is glossed as ‘who?’.

Mongo

- (21) *lina likě ná/*é?*
 G5.name AG5.2SG.POSS who?/*what?
 ‘What is your name?’ (Honoré Vinck, p.c.)

2.1.1.2.3 Possible origins of the lack of differentiation between ‘who?’ and ‘what?’ in Mongo (and some other languages of zone C)

The Mongo interrogative *é* ‘what?’ clearly goes back to the Proto Bantu interrogative pronominal root **í*, which on a solid comparative evidence can be reconstructed with the meaning ‘what?’ (or, in combination with one of the three Proto Bantu locative gender prefixes, ‘where?’ and ‘which one?’; cf. Meeussen

1967:103; *Bantu Lexical Reconstructions 3*;¹⁴ Doneux 1971:131-138; Doneux & Grégoire 1977:186-187). The situation with the Mongo interrogative *ná* ‘who?, what?’ is less clear. Judging from Meeussen (1967:103), Doneux & Grégoire (1977:193), and the database *Bantu Lexical Reconstructions 3*, *ná* and similar forms in other Bantu languages can be brought back to the following forms: **n(d)á-í*, **ndá-ní* or **(n)á-ní*, neither of which is reconstructable (with certainty) to the Proto Bantu level. The three aforementioned sources suggest to reconstruct these forms with the meaning ‘who?’. However, this semantic reconstruction makes it very difficult to account for the presumed shift from ‘who?’ to ‘who?, what?’ in Mongo and several other languages of zone C.

I believe that the interrogatives **n(d)á-í*, **ndá-ní* and **(n)á-ní* should rather be reconstructed as selective interrogatives meaning ‘which one? (person or thing)’, and ultimately as locative ‘where?’. This reconstruction would square well with both (i) the fact that in many languages the reflexes mean ‘who?’¹⁵ and (ii) the fact that in a few languages, such as Mongo, the reflexes can be used as both ‘who?’ and ‘what?’. The latter development can be explained as an expansion of the selective ‘which one? (person or thing)’ to non-selective contexts, with its original indifference to the distinction between the features person and thing simply being maintained. This hypothesis can also help us to account in a straightforward way for (i) the attributive use (with both human and non-human nouns) of the form that also means ‘who?’, and for (ii) the development of ‘who?’-dominance in NAME-questions, as was presented in Section III.2.1.1.2.2 and can be illustrated here with example (22a) from Ngombe, another language of zone C.

Ngombe (Bantu C50; Democratic Republic of the Congo; Rood 1958:xxi)¹⁶

- (22) a. *ngando íyě nda?*
 G9.village AG9.this who
 ‘What is the name of this village? (lit.: ‘Who is this village?’)
- b. *moto íyě nda?*
 G1.man AG1.this who
 ‘Who is this man?’

Thus, as far as NAME-questions are concerned, as discussed Section II.3.1, it is

¹⁴ Available at: <http://www.metafro.be/blr>.

¹⁵ In this respect, recall that there exists a common semantic link between ‘which one?’ and ‘who?’ (cf. Section I.2.3). This link can also be found, for instance, in many Bantu languages of zone J, where according to Doneux & Grégoire (1977:193), ‘who?’ has been replaced by the form originally meaning ‘which one?’.

¹⁶ The variety of Ngombe described by Rood (1958) is Ngenja, or Ligenza in Ethnologue’s denomination.

not uncommon cross-linguistically for ‘where?’ and especially ‘which one?’ to be used as avoidance strategies in questions about proper names. If we suppose that this was also the case in the Bantu languages at issue, this would mean that when the shift from (‘where?’ and) ‘which one?’ to ‘who?’ occurred, many of these languages must have simply kept on using the same interrogative in questions about proper names (cf. the discussion of an apparently similar evolution in Cushitic languages in Section II.3.3.1).

The locative-selective hypothesis proposed for the interrogatives **n(d)á-í*, **ndá-ní* and **(n)á-ní* is also plausible from the formal point of view. Thus, Doneux & Grégoire (1977:188) remark that in most languages of zone C the stem of the locative interrogative is an innovation as compared to Proto Bantu *pa-í* ‘where?’ (< *pa-* G16, which is a locative gender marker, and the stem *-í* ‘what?’). Furthermore, Doneux & Grégoire (1977:186-188) provide quite a few examples of languages from zones A, B, H, K and L where locative interrogatives are based on the root *-ní* rather than *-í*, as e.g. in Tsogo *va-ní* (Bantu B30; Gabon) or Duma *ya-ní* (Bantu B50; Gabon). Finally, in Mongo and quite a few other languages there exist a general locative preposition *n(d)á* (and the like), as in Mongo *ndá ilombe* ‘in the house’, *ndá loulu* ‘on the roof’ or *ndá lisála* ‘(he went) to the field’ (Hulstaert 1938:134). All this seems to suggest that (i) at some earlier stage in the languages of zone C (and perhaps some others as well) the Proto Bantu interrogative *pa-í* ‘where?, which one? (person or thing)’ was first replaced with **n(d)á-í* (or **ndá-ní*, **(n)á-ní*) ‘where?, which one? (person or thing)’ (in both cases, the locative meaning is clearly the original one),¹⁷ (ii) subsequently, **n(d)á-í* (or **ndá-ní*, **(n)á-ní*) ‘where?, which one? (person or thing)’ became restricted to pronominal use as ‘which one?’, with a parallel evolution of a new locative interrogative, and (iii) finally, the selective ‘which one? (person or thing)’ was expanded to non-selective contexts, in many languages only as ‘who?’ but in some, such as Mongo, as both ‘who?’ and ‘what?’.

2.1.2 Atlantic languages: Bijogo

Bijogo/Bidyogo/Bijago is a small language spoken on the Bijagós Archipelago in Guinea-Bissau. It forms on its own a separate branch within the Atlantic family of the Niger-Congo phylum. Segerer (2002) describes the variety of Bijogo spoken in the village Bijante on Bubaque Island. This variety belongs to the central group of dialects, known as Kagbaaga. Bijogo appears to be the only Atlantic language allowing for a lack of differentiation between ‘who?’ and ‘what?’.

¹⁷ There might have also been an intermediate stage on which Proto Bantu *pa-í* was replaced by *pa-ní*.

Bijogo has two interrogative pronominals, viz. *we/w-* meaning ‘who?’, as in (23), and *ɲɔ* (after a vowel)/-ɔ (after a consonant), which normally means ‘what?’, as in (24).

Bijogo

- (23) a. [A:] *ɲɔ-nam* *we?* [B:] *ɲɔ-nam* *Pedro*
 AG<ɲɔ>.PFV-be IPW AG<ɲɔ>.PFV-be PROP
 ‘[A:] Who is it? [B:] It’s Pedro’ (Segerer 2002:150)
- b. *w-ɔn-tukp-ak* *Mario?*
 IPW-AG<ɔ>.PFV.FOC-hit-PFV PROP
 ‘Who hit Mario?’ (Segerer 2002:266)
- (24) a. *mi-ní* *ɲɔ?*
 2SG.IPFV-drink IPW
 ‘What do you drink?’ (Segerer 2002:81)
- b. *mi-rór-ɔ?*
 2SG.IPFV-look.for-IPW
 ‘What are you looking for?’ (Segerer 2002:81)
- c. *ɲun-nam-ɔ?*
 AG<ɲɔ>.PFV.FOC-be-IPW
 ‘What is this?’ (Guillaume Segerer, p.c.)

However, according to Guillaume Segerer (p.c.), *ɲɔ/-ɔ* can also mean ‘who?’ when used predicatively with the copula *nam* ‘be’, as in (25), which is “the formula generally used for asking the name of someone”. Compare (25) with (24c).

Bijogo

- (25) *an-nam-ɔ?*
 2SG.PFV.FOC-be-IPW
 ‘Who are you?’, ‘What’s your name?’ (Guillaume Segerer, p.c.)

As pointed out by Guillaume Segerer (p.c.), the form *ɲɔ* serves not only as the interrogative ‘what?’ but also as the “generic/neutral” gender and agreement pattern prefix, as in (23a) above. This prefix is also found in the word ‘thing’ *ɲo-o* (< *ɲɔ-o*). It is not implausible that the non-human interrogative and the ‘thing’-prefix are related historically. It may be interesting to note in this respect that the postconsonantal allomorph of the interrogative ‘what?’, viz. *ɔ*, which in (25) also means ‘who?’, is identical to the singular human gender and agreement pattern prefix *ɔ-*, as in (23b) above or in words like *ɔ-kanto* ‘woman’ and *ɔ-gbaga* ‘inhabitant of the island Bubaque’ (Segerer 2002:101). Therefore, we may hypothesize that the formulaic expression in (25) represents a frozen relic of a previous stage of the language when *ɔ* meant ‘who?’. That *ɔ* is no more widely

used in the meaning ‘who?’ in the modern language may primarily be accounted for by a later development of the postconsonantal allomorph of *ɲɔ* ‘what?’, viz. *ɔ*. Such a development must have resulted in a conflict between the two meanings of *ɔ*, which has been resolved in favour of the non-human meaning, except in cases like (25) where the human interpretation is made obvious by the semantics of the subject marker.

2.2 “Khoisan” languages

Two “Khoisan” languages have been found to use a single form for ‘who?’ and ‘what?’: Eastern !Xóǝ (Lone Tree !Xóǝ) and, probably, #Hǝǝ. The structural similarity is quite interesting because although the two languages are spoken in the immediate vicinity of each other, they appear to be genetically unrelated according to the most recent studies (see, for instance, Güldemann & Vossen 2000, Güldemann forthcoming). This is why I put the word *Khoisan* in inverted commas. However, the latest edition of the Ethnologue (Gordon 2005) classifies !Xóǝ and #Hua (as #Hǝǝ appears to be labelled there) as two closely related languages constituting the Hua group of the Southern branch of Southern African Khoisan.¹⁸ Therefore, I will begin by considering the issue of genetic classification in more detail in Section III.2.2.1. In Section III.2.2.2, I will present the !Xóǝ data and in Section III.2.2.3 the #Hǝǝ data.

2.2.1 Genetic classification of “Khoisan” languages

In what follows, I base myself on Sands (1998), Güldemann & Vossen (2000) and, especially, Güldemann (forthcoming).

The term *Khoisan* summarily refers to all East and South African languages with clicks, other than Bantu and Cushitic. Since Greenberg’s (1966b) genealogical classification of the languages of Africa, Khoisan languages are widely believed to constitute one linguistic phylum.¹⁹ However, genetic unity of Khoisan languages is one of the most controversial parts of Greenberg’s classification and linguists working on Khoisan languages have been criticizing it right from the beginning (cf. Westphal 1956, 1962a, 1962b). Nowadays, there seems to be a consensus among Khoisanists that what is usually referred to as Khoisan languages represent at least four (and perhaps even seven) different

¹⁸ It is not excluded, though, that Ethnologue’s #Hua refers to the #Hǝǝ variety of Eastern !Xóǝ (the latter is mentioned in Güldemann 2005:6). If this is the case, then #Hǝǝ has not been taken up into Ethnologue at all. Otherwise, Ethnologue must have lumped #Hǝǝ together with a variety of !Xóǝ.

¹⁹ The first published versions of Greenberg’s classification of African languages go back to 1949. Greenberg (1966b) is the final, more widely known version.

lineages. There is no sufficient evidence for positing genetic relationship between them. Güldemann (forthcoming) speaks about six lineages, as presented in Figure 1, which are potentially reducible to four lineages. Sandawe may be distantly related to Khoe-Kwadi languages and †Hõã should probably be grouped together with the Ju languages. Güldemann proposes to divide the 6 lineages into three “pragmatically oriented groups”, as in (26).

*Figure 1. Lineages subsumed under Khoisan and their internal composition
(based on Güldemann forthcoming)*

Hadza

Sandawe

Khoe-Kwadi

Kwadi

Khoe (= Central Khoisan)

Khoekhoe (Hottentot)

North: Eini[†], Nama-Damara (Nama), Hai||'om

South: !Ora[†] (Korana), Cape varieties[†]

Kalahari

East

Shua: Cara, Deti[†], |Xaise, Danisi, Ts'ixa, etc.

Tshwa: Kua, Cua (Hiechware), Tsua, etc.

West

Kxoe: Kxoe, ||Ani, Buga, G|anda, etc.

G||ana: G||ana, G|ui, †Haba, etc.

Naro: Naro, etc.

Ju (= Northern Khoisan)

Northwest: !'O!Xũu, !Xũu (!Kung)

Southeast: Ju|'hoan, †Kx'au||'e (Auen)

†Hõã (†Hõã, †Hua)

Tuu (= Southern Khoisan)

Taa: !Xõo_{DC} (!Xóõ, !Xõõ), N|amani[†], N|u||'en[†] (|Nu||en = Nusan),

Kakia[†]_{DC} (Masarwa)

Lower Nossob: |'Auni[†] (|Auni), |Haasi[†]_{DC}

!Ui: N||ng, †Khomani, N|huki_{DC}; †Ungkue[†]; |Xam[†]_{DC} (|Kam); ||Xegwi[†]

[†] = extinct, _{DC} = dialect cluster

- (26) A “pragmatically oriented” division of the six “Khoisan” lineages (based on Güldemann forthcoming)
- (A) Hadza and Sandawe, the two East African Khoisan languages that show little evidence of being related to other Khoisan languages
 - (B) the genealogical Khoe-Kwadi group in Southern Africa
 - (C) non-Khoe: typologically similar but genetically unrelated Ju and Tuu families of Southern Africa, together with one unclassified language, †Hõa, which may be related to the Ju family

2.2.2 Eastern !Xóõ (“Khoisan”, Tuu, Taa)

As noted by Güldemann (2005:4-5), interrogative proforms in Tuu languages are “frequently rendered by the co-occurrence of a general question marker with an indefinite proform which conveys the onomasiological category of the referent under consideration [...] The indefinite proform can be a generic noun, a pronoun, or even a verbal item”. The latter case is represented in Eastern !Xóõ by the locative predicate *āh’ā* ‘be somewhere’, which in questions can be translated, depending on the context, either as ‘(be) where?’, as in (27), or as ‘(be) which?’, as in (28). The use of a generic noun is illustrated in (29) from N|huki with the word *tyú* ‘person’.

Eastern !Xóõ (Tuu, Taa; Botswana; Traill 1994:18)

- (27) /é Bōlo //Xáo ní āh’ā /núm tshúu /ĩi
 Q-AG3 PROP[G3] TENSE be.somewhere stay sit be
 ‘Where does *Bōlo* //Xáo live?’²⁰
- (28) /-ā ní bà káne ká ’āā-sa tǎ
 Q-2SG TENSE IPFV want MPO.AG2 eat-NMLZ[G2] REL.AG2
āh’ā kǎ?
 be.somewhere REL.AG2
 ‘Which food do you want? (lit.: ‘You want the food that is where?’)’

N|huki (“Khoisan”, Tuu, !Ui; South Africa)

- (29) *tyú xè ’à Ōwà?*
 person Q 2SG child
 ‘Who is your child?’ (Westphal n.d. cited via Güldemann 2005:16)

Generic nouns meaning ‘person’ and ‘thing’ are the most common kind of “indefinite proforms” used in Tuu languages to form interrogative pronominals

²⁰ The glosses of Eastern !Xóõ examples are based on Traill (1994) and Güldemann (2005). Genders and agreement patterns are numbered following Traill (1994) for Eastern !Xóõ and following Güldemann (2006) for Western !Xóõ.

meaning ‘who?’ and ‘what?’, respectively. A locative predicate is found regularly for the meanings ‘where?’ and ‘which?’, although for ‘which?’ in some languages this holds only historically (cf. Güldemann 2005:16). Eastern !Xóǝ is special in that it uses a third person pronoun to form an interrogative pronominal and in that it has only one interrogative pronominal which does not differentiate between ‘who?’ and ‘what?’. In fact, the latter peculiarity is related to the former because Eastern !Xóǝ uses one and the same pronoun *èh* of agreement pattern 3 for both meanings, as illustrated in (30).

Eastern !Xóǝ (Traill 1994:18)

- (30) */-ā ní bà káne ké èh?*
 Q-2SG TENSE IPFV want MPO.AG3 PRN.AG3
 ‘Whom/ what do you want?’

In order to explain the lack of differentiation between ‘who?’ and ‘what?’ in Eastern !Xóǝ, a few words should be first said on the gender and agreement system of this language, as presented in Traill (1994). Gender in Eastern !Xóǝ is mostly “covert”: nouns do not carry any gender marking morphology and gender manifests itself only on agreement targets. There are five agreement patterns in Eastern !Xóǝ. Genders do not have any clear semantics and often they are not number-sensitive either. Gender assignment is largely unpredictable. Similar systems are found in other Tuu languages and in the Ju family (Güldemann & Vossen 2000:111, Güldemann 2005:17). The number of genders and the kinds of agreement targets that have gender marking may differ but minimally agreement will be expressed with pronouns. Other Khoisan lineages have a radically different gender system, which is primarily sex-based distinguishing masculine and feminine genders, as well as a third so-called “common” gender in the case of Khoe languages (Güldemann & Vossen 2000:111-120). †Hǝǝ appears to be the only lineage lacking genders.²¹

Although gender assignment is largely unpredictable in Eastern !Xóǝ, certain patterns may be discerned. For instance, there is “a fairly consistent semantic association” between gender/agreement pattern 3 and nouns denoting “singular living beings” (especially humans) and gender/agreement pattern 4 and nouns denoting “plural living beings” (especially humans), although there are “many exceptions” as well (Traill 1994:22). This trend is particularly prominent in the case of pronouns (Traill 1994:22). An explanation for this trend may lie in the fact that the generic nouns *tāa* ‘person’ and *tùu* ‘people’ belong to genders 3 and 4 respectively. This would also explain why the pronoun *èh* of agreement pattern 3 has been chosen as a basis for the interrogative pronominal construction

²¹ Cf. <http://instruct1.cit.cornell.edu/courses/ling700/hoan.htm>.

meaning ‘who?’. The lack of a corresponding plural form of ‘who?’ (as well as of ‘what?’) is not particularly surprising from a typological perspective.

It is somewhat more difficult to explain why the same pronoun *èh* has been chosen as a basis for the interrogative pronominal construction meaning ‘what?’, for the generic noun *tháa* ‘thing’ is of gender 5. Therefore, the explanation should be sought elsewhere and I believe that the following facts are of special interest here. First, “the vast majority of foreign acquisitions” are assigned agreement pattern 3 in the singular and agreement pattern 4 in the plural “no matter what their phonological shape” and meaning are (Traill 1994:22). Furthermore, it seems that agreement pattern 3 can be used for the purposes of so-called enforced agreement (cf. Section III.2.1.1.2.1), e.g. when no controller is present at all.²² This is illustrated in (31), where the pronoun and the demonstrative do not agree with the gender 1 word *g!xá’u* ‘south wind’.

Eastern !Xóǀ

- (31) *èh té’è ñ k! g!xá’u*
 PRN.AG3 this.AG3 TENSE COP.AG1 south.wind[G1]
 ‘This one is the south wind’ (Traill 1994:87)

These two facts indicate at agreement pattern 3 (and the respective gender) as the default option in Eastern !Xóǀ, which may explain its choice as a base for the interrogative pronominal construction meaning ‘what?’. Finally, it is not inconceivable that the choice of *èh* (phonologically, /ǝ/) for ‘what?’ might have also been influenced by similarly sounding interrogative elements in neighbouring languages. For instance, one may think about the question marker *ǝ* found in many Kalahari Khoe languages (see examples in Vossen 1997:264-267) or the ‘what?’ interrogatives in the neighbouring Bantu languages, such as *eng* ‘what?’ in Tswana (Bantu S30; Botswana, Namibia, South Africa), Botswana’s national language and lingua franca.

Interestingly, proper names also seem to trigger agreement pattern 3, as can be seen in (27), where the question marker agrees with the proper name subject. Although this is not directly relevant for the question at issue, it happens to fit very well in the overall picture of the use of agreement pattern 3 in Eastern !Xóǀ. The association of human nouns, loan words, enforced agreement contexts, and proper names with one and the same agreement pattern is strongly reminiscent of what one finds in many Bantu languages, where one and the same “human” agreement pattern 1 is also often used in exactly the same situations (cf. Section

²² As far as I can judge, another pattern commonly used for this purpose is agreement pattern 5 corresponding to the agreement pattern of the word *tháa* ‘thing’. Typologically, this would seem to be a more natural choice than agreement pattern 3, since, as Corbett (1991:206) notes, agreement patterns linked to human genders are generally avoided in such cases.

III.2.1.1.2.1; Van de Velde 2006). One may even speculate that this is due to Bantu influence, for in the closely related Western !Xóǝ the situation is somewhat different.

In Western !Xóǝ, pronouns of agreement patterns 3 and 4 in singular and plural, respectively, tend to be used with “human/ animates irrespectively of their lexical gender” (Güldemann 2006), as in (32). In this respect, Western !Xóǝ behaves similarly to Eastern !Xóǝ. However, proper names always trigger agreement pattern 1 (Güldemann 2006), as in (33).

Western !Xóǝ (Güldemann 2006)

- (32) */huun i ʔqhaa n/nn [...] e ʔqhaa*
 white.man[G1] PRN.AG1 then say PRN.AG3 then
g//u’ri lori ʔose
 open.AG1 lorry[G1] door[G?]
 ‘The Boer said, [...] he then opened the lorry’
- (33) *nna n/ai Tom-tu ku /ai*
 1SG.PRF see.AG1 PROP[G1]-ASS.PL[G4] REL.AG4 stay
ki Dertien ku
 MPO.AG1 PROP[G1] REL.AG4
 ‘I have seen Tom and them who were at Post 13’

Interestingly, loans in Western !Xóǝ seem to trigger the same agreement pattern as proper names, which is reminiscent of Eastern !Xóǝ. At least, all clear loans in Güldemann (2006) do so: *lori* ‘lorry’, *tuin* ‘garden’ (< Afrikaans *tuin*) and *seleng* ‘money’ (< *shilling* via Bantu languages) all trigger agreement pattern 1. Information on the agreement pattern used for enforced agreement in Western !Xóǝ is lacking (Tom Güldemann, p.c.).

Western !Xóǝ also differs considerably from Eastern !Xóǝ in the way it builds interrogative pronominals meaning ‘who?’ and ‘what?’. The interrogative ‘who?’ is realized by means of a relative construction consisting of the locative (predicate?) *haan* ‘(be) where?’ (clearly cognate to the Eastern !Xóǝ *āh’ā* considered above), which here can be translated as ‘(be) which?’, followed by a relative marker of agreement pattern 3 or 4 when the referent is supposed to be singular or plural, respectively. This is illustrated in (34), which can be literally translated as something like *You see the one(s) that is(are) where/which?*²³

²³ In construing ‘who?’ as a combination of ‘which?’ and an agreeing element, Western !Xóǝ very much resembles neighbouring Khoe languages. In these languages, as described by Vossen (1997:260-265), ‘who?’ and ‘which?’ are typically expressed by one form, which can be optionally marked for gender (it is “konkordanzbereit”, in Vossen’s terms). For proto-Khoe Vossen (1997:379-381) reconstructs this form as **mā* and of the two meanings, ‘who?’ and ‘which?’, he arbitrarily chooses ‘who?’ as the original meaning. However, both semantically

Western !Xóǎ

- (34) a. *a si n/a-e haan ke?*
 2SG IPFV see-AG3 where REL.AG3
 ‘Who do you see?’ (DoBeS-team via Tom Güldemann, p.c.)
- b. *a si n/a-u haan ku?*
 2SG IPFV see-AG4 where REL.AG4
 ‘Who(PL) do you see?’ (DoBeS-team via Tom Güldemann, p.c.)

Compare also (34) to (35), where *haan* is used in its original meaning.

Western !Xóǎ

- (35) *a haan si n/ann /a †xanya*
 2SG where IPFV see.1SG GEN.AG2a book[G2b]
 ‘Where do you see my book?’ (Güldemann 2006)

Unlike ‘who?’, the interrogative ‘what?’ in Western !Xóǎ is a nominal, *thi*, which triggers agreement pattern 2a, as illustrated in (36).²⁴

Western !Xóǎ

- (36) *a si n/a-an thi?*
 2SG IPFV see-AG2a IPW[G2a]
 ‘What do you see?’ (DoBeS-team via Tom Güldemann, p.c.)

2.2.3 †Hǎǎ (“Khoisan”, isolate or Ju)

There is extremely little information available on †Hǎǎ, most of which is in form of unpublished field notes. To my knowledge, the only “published” source on †Hǎǎ that gives any information on interrogative proforms is the †Hǎǎ part of the website on the syntax of Khoisan languages authored by Chris Collins and colleagues at Cornell University (<http://instruct1.cit.cornell.edu/courses/ling700/hoan.htm>), where I found the following examples.

†Hǎǎ (Isolate or Ju; Botswana)

- (37) *’ar’’i ya †’u ’a-ci?*
 IPW Q duiker PROG-see
 ‘Who does the duiker see?’

and typologically, the meaning ‘which?’ would have been a much more plausible choice than ‘who?’ (cf. Section I.2.3).

²⁴ Agreement patterns 2a and 2b differ only by their tones, low /-ǎ/ and high /-ǎ́/, respectively (Güldemann 2006). However, tone marking is lacking in the Western !Xóǎ examples cited in the present Section because they are not marked in the sources.

- (38) O'u ki-//kain ya u 'a-ci?
 duiker IPW Q 2SG PROG-see
 ‘Which duiker do you see?’
- (39) u 'a-'am //ka''e ya?
 2SG PROG-eat meat Q
 ‘Are you eating meat?’

If the gloss in (37) is correct, then #Hõã will be the second “Khoisan” language next to Eastern !Xóõ that uses a single form for ‘who?’ and ‘what?’. The fact the two languages are spoken in the immediate vicinity of each other provides an indirect indication that this is indeed the case. Furthermore, an unpublished wordlist in Gruber (1975) gives two forms both glossed as ‘who?, what?’: 'ã''ri with a variant n!nã''ri .²⁵ Unfortunately, my attempts to contact Chris Collins and Jeff Gruber to clarify this issue have not had any success.

2.3 Afro-Asiatic languages

2.3.1 Introduction

The Afro-Asiatic phylum consists of at least six major branches: Berber, Egyptian, Semitic, Chadic, Cushitic and Omotic. Languages with interrogative pronominals used indiscriminately as ‘who?’ or ‘what?’ appear to exist in four of them, Berber (Section III.2.3.2), Egyptian (Section III.2.3.3), Semitic (Section III.2.3.4), and Cushitic (Section III.2.3.5).

The branches differ in how numerous the languages with ‘who?, what?’ interrogatives are. While they seem to abound in Berber, Semitic has to be content with just a few, and Cushitic has only one. Egyptian is a single-language branch. Egyptian data must be considered with some caution, though, since it is an extinct language with much of the data recorded in a less than perfect orthography. As to the remaining two branches, Chadic and Omotic, they seem to lack such ‘who?, what?’ interrogatives. Note, however, that they are also the least described branches and Chadic is at the same time by far the biggest branch as regards the number of its living members.

The branches also appear to differ in the ways they have arrived at the lack of formal differentiation between ‘who?’ and ‘what?’. In Berber, where existence of ‘who?, what?’ interrogatives seems to be particularly ancient, languages also have surprisingly large inventories of interrogative pronominals. In all probability, the two facts are related and are both primarily due to the fact that Berber interrogative pronominals are not single morphemes but constructions consisting of a general interrogative root ‘who?, what?, which one?’ (normally,

²⁵ The sign <’’> marks pharyngealization of the vowel, the rising tone is marked only on the first vowel.

m) and a demonstrative pronominal that typically introduces a relative clause. Structurally, these constructions are similar to those used for the purpose of focalization and can be analyzed as clefting. The kind of demonstrative pronominal used determines whether and how a distinction between ‘who?’ and ‘what?’ is made (cf. Sections III.2.3.2.2.1 and III.2.3.2.3).

Provided the analysis in Section III.2.3.3.2.1 is correct, Egyptian must have also had a general interrogative root ‘who?, what?, which one?’ of the form **/m/* or **/mi ~ ma/*, which gradually disappeared as an independent entity in later stages of Egyptian’s history. For a certain time, a new interrogative ‘who?, what?, which one?’ has developed from the masculine singular demonstratives *pw*, but it has been lost later as well (Section III.2.3.3.2.2). Finally, with one interrogative, viz. *zy*, Egyptian also appears to have allowed a lack of differentiation between ‘who?’ and ‘what?’ in predicative use (Section III.2.3.3.2.4).

In Semitic, in the cases where enough unambiguous data are available, ‘who?, what?’ interrogatives seem to be relatively recent semantic expansions of the interrogative **man* (or the like) ‘who?’. In all probability, these developments have been instigated by a preceding expansion to the meaning ‘who?’ of the interrogative root **ʔay* (or the like), originally ‘where?’, but also pronominally in the meaning ‘which one?’ and/or attributively as ‘which [N]?’, what (kind of) [N]?’. The expansion of **ʔay* (or the like) to the meaning ‘which one?’ must have resulted in a certain competition between the two interrogatives, which apparently involved some oscillation in the semantics of the original interrogative ‘who?’. Furthermore, with the “Canaanite-Akkadian” mixed language, an extinct mixed East-West Semitic idiom (ca. 1350 BC), Semitic may provide an interesting example of the development of a ‘who?, what?’ interrogative as a result of an imperfect fusion of the superstrate with the substrate, East and West Semitic languages respectively, which happened to have the same form, or two very similar forms, but with opposite meanings, ‘who?’ and ‘what?’. Finally, although most of the attested cases of ‘who?, what?’ interrogatives in Semitic are likely to be relatively recent, a cross-Semitic comparison of interrogative pronominals suggests (cf. Section III.2.3.4.1.1) that Proto-Semitic may have had at least one interrogative pronominal ‘who?, what?’, probably **mi*, next to **man* ‘who?’ and **ma* ‘what?’. However, it must have become specialized in one of the two meanings rather early in the history of Semitic languages through its fusion with a deictic element that followed it, that is due to conventionalization of a Berber-like structure [interrogative + deictic], which is probably best analyzed as clefting.

In the single Cushitic language with a ‘who?, what?’ interrogative, Saho, this interrogative must have evolved somewhat similarly to what we find in Semitic. The main difference is that it is the interrogative root **ʔay* (or the like),

originally ‘where?’, but also pronominally in the meaning ‘which one?’ (and/or attributively as ‘which [N]?, what (kind of) [N]?’), that has prevailed after having been expanded to cover non-selective contexts as well, in questions about both things and persons.

The fact that languages with ‘who?, what?’ interrogative pronominals appear to exist in four of the six major branches of Afro-Asiatic considered together with the phylum’s immense age, ranging from 12000 to 20000-30000 years, depending on the scholar (cf. Militarev 2005:340, 398), may be indicative of some kind of structural inclination going back to the early stages in the history of Afro-Asiatic. More recent innovations of ‘who?, what?’ interrogatives provide further support for the existence of such a structural inclination. I believe that this inclination must be due to the way the system of interrogative pronominals was organized in early Afro-Asiatic languages. To various extents, traces of this system appear to persist in the more recent languages. I have tried to summarize those that seem to be the most relevant in (40).

(40) *Some relevant features of interrogative pronominal systems in Afro-Asiatic languages*

- a. A relatively high number of languages with interrogative pronominals used indiscriminately as ‘who?’ or ‘what?’.
- b. The most recurrent forms that interrogative pronominals show in the Afro-Asiatic languages can be organized in two groups, non-selective **m ~ mi ~ ma* ‘who?, what?’, and selective **ʔay* (or the like) ‘which one?’. The latter clearly goes back to an interrogative ‘where?’ (cf. Lipiński 1997:328).
- c. There seem to be rather many languages that do not allow, or at least disprefer, attributive constructions of the ‘which [N]?, what (kind of) [N]?’ type. To a large extent, this may be due to the ‘where?’ origin of the interrogatives with such semantics.
- d. Widespread tendency to conventionalize constituent questions as clefts of the structure [interrogative + deictic]

Feature (40a) suggests that in some proto-stage such ‘who?, what?’ interrogatives may have existed as well. As for (40b), many forms of the interrogative pronominals can be found further in sections on particular branches of Afro-Asiatic, but to give some support to my statement here and now, let us simply consider some of the available reconstructions of Afro-Asiatic interrogative pronominals. Note that the reconstructions hardly ever mention the selective interrogative ‘which one?’. In his overview of the Afro-Asiatic phylum, Diakonoff (1965:75 [1988]) mentions a pronominal base *m-*, which has served as the source for “interrogative (and indefinite) pronouns” in many Afro-Asiatic

languages, and which was in all probability originally “demonstrative in nature (for reference to distant and distant invisible objects)”. Sasse (1981:143) reconstructs the interrogative pronominals **ʔay* and **ʔaw* for animate entities (“belebte Entitäten”), and **wa-* and **ma-* for inanimate entities (“unbelebte Entitäten”). Bender (1990:667) gives Afro-Asiatic **ma(n,t)-* ‘what?’ and **mi(n,t)-* ‘who?’, and in (2000:229), he mentions an Afro-Asiatic “interrogative” **ay*. Ehret (1995:300-301, 453) reconstructs **m-* as “indefinite pronoun stem (‘one, someone, somebody’)”, and both **ma ~ *mi* and **wa ~ *wi* as ‘what?’, even though ‘who?’ is a rather common meaning of their alleged reflexes as well.

I believe that the most plausible conclusion suggested by the available data would be to reconstruct **m-i ~ *m-a* as a ‘who?, what?’ interrogative and **ʔay* as originally ‘where?’, but also pronominally in the meaning ‘which one?’.²⁶ It is most likely that the two forms **m-i* and **m-a* did not differ lexically, as ‘who?’ vs. ‘what?’, at least not directly so. Rather the two vowels marked a difference in case, e.g. *-i* for a kind of genitive and/or nominative (ergative) and *-a* for accusative (absolutive) and predicate nominals.²⁷ It is not uncommon for Afro-Asiatic languages to formally group these functions together in a comparable way. For instance, the nominal morphology of Berber languages typically makes a formal distinction between the so-called “annexed state”, used for marking possessors, nouns headed by an adposition, and regular (non-topicalized) subjects among other things, and the so-called “free state”, used for marking direct objects, predicate nominals, citation forms, and topicalized subjects among other things (cf. Aikhenvald 1995a:44-45). Considering the usual correlations between animacy/referentiality and case, it would be natural to expect a closer link between the genitive/nominative/ergative and human ‘who?’, on the one hand, and between the accusative/predicative and a non-human ‘what?’, on the other (cf. Nau 1999). Indeed, there appears to be a certain tendency in Afro-Asiatic languages for the reflexes of the presumably “accusative” form **ma* to specialize as ‘what?’ and for the reflexes of the presumably “genitive” form **mi* to specialize as ‘who?’ (cf., e.g., Sections III.2.3.2.3 and III.2.3.4.1.1). A few noticeable deviations among the reflexes of **mi* are probably due to the fact that the “genitive” is the case regularly used to mark nominals headed by an adposition, and adpositional phrases are typically used to code non-core

²⁶ Other interrogatives might have existed as well, or it might have been possible to use nouns with generic meanings such as ‘person’ or ‘thing’ to form constituent questions. However, even if all these extra possibilities really existed, this would have little impact on the point made here.

²⁷ For the forms of the case markers see Bender (1990:666, 2000:210-212), Diakonoff (1965:54-57 [1988]), and Lipiński (1997:253-265) among others. In principle, the difference between **mi* and **ma* could have also been that of gender, masculine vs. feminine, the only two genders that Afro-Asiatic languages normally have in the singular. However, the available reconstructions do not seem to provide enough support for the opposition **-i* vs. **-a* as masculine vs. feminine.

arguments, with all the natural implications with respect to animacy this involves. Possible reasons for deviations among the reflexes of **ma* will be considered below.

The feature (40c) can be illustrated with data from Ahaggar Tuareg (Southern Berber, North Tuareg; Algeria), which may have attributive constructions of the ‘which [N]?, what (kind of) [N]?’ type only within a couple of fixed expressions, e.g. *man-ăket* ‘how much?, how many?, how long (time)? (lit. ‘what measure’))’ (Karl-G. Prasse, p.c.), but the attributive interpretation is not incontestable (see Section III.2.3.2.2.4). Elsewhere a cleft construction of the kind illustrated in (41) and (42) needs to be used.

Ahaggar Tuareg (Southern Berber, North Tuareg; Algeria)

- (41) *manek-k aləs wa tənăyăd ənğom?*
 IPW-3SG.M.OBJ man DEM.M.SG 2SG.SBJ.saw previously
 ‘Which man (among those present now) did you saw previously? (lit.: ‘Which is he (among those present now), the man you saw previously?’)’
 (Karl-G. Prasse, p.c.)
- (42) *mī yămos aləs wa tənăyăd ənğom?*
 IPW 3SG.M.SBJ.be man DEM.M.SG 2SG.SBJ.saw previously
 ‘Who is the man (among those present now) which you saw previously?’
 (Karl-G. Prasse, p.c.)

As to (40d), ample illustrations can be found e.g. in Sections III.2.3.2.2 and III.2.3.4.1.2.

In a way, the fact that the attributive use of interrogatives is dispreferred may favour the blurring of the differentiation between selective and non-selective interrogatives. Presumably, the reason is that when selective and non-selective interrogatives always have the same syntactic function, keeping up the semantic distinction between them may become somewhat more difficult. Ultimately, this may result in a lack of formal differentiation between the selective and non-selective interrogatives, with the selective interrogative expanding to cover the non-selective contexts. This may explain the large-scale expansion of the originally selective interrogative ‘which one?’ to non-selective meanings ‘who?’ or ‘what?’ in Cushitic and Omotic (cf. Section III.2.3.5.1). Furthermore, the tendency to conventionalize constituent questions as clefts, taken together with the fact that predicate nominals are typically marked in the same way as direct objects, may favour generalization of the reflexes of the originally “accusative” form **ma*.

Depending on the kind of deictic used, the preferred structure of the cleft, [interrogative + deictic], may either help to create/keep a distinction between ‘who?’ and ‘what?’ or help to keep it unexpressed. The variation among the Berber languages (Section III.2.3.2) is particularly interesting in this respect.

2.3.2 Berber languages²⁸

Languages with interrogative pronominals used indiscriminately as ‘who?’ or ‘what?’ can be found in all extant major branches of Berber languages, probably with the exception of the Eastern branch. Here I use the classification proposed by Aikhenvald (1984) and Aikhenvald & Militarev (1991), which distinguishes four living branches, presented in Figure 2, next to a few extinct ones. As Aikhenvald & Militarev (1991:157) point out, lower level classifications of the more than two hundred Berber idioms are still largely based on geographical criteria due to the lack of necessary descriptions. The classification of Berber found in the Ethnologue is quite similar to that of Aikhenvald & Militarev, at least as far as the major division in four living branches is concerned. It should be mentioned, however, that according to Maarten Kossmann (p.c.), Aikhenvald & Militarev’s classification is hardly ever used by most linguists working on Berber languages. As I am not in position to make any educated judgements on this issue, my use of Aikhenvald & Militarev’s classification here should be viewed as being due to purely practical considerations.

I will not attempt to give an exhaustive list of Berber languages with a ‘who?, what?’ interrogative. Instead, in Section III.2.3.2.1 I will indicate at least one such language from a branch. In Section III.2.3.2.2, I will discuss some important structural peculiarities of Berber constituent questions, such as their typically constructional nature. These peculiarities result in a multitude of forms of interrogative pronominals in Berber languages and several ‘who?, what?’ interrogatives. In Section III.2.3.2.3, a brief summary will be given.

2.3.2.1 Some Berber languages with ‘who?, what?’ interrogatives

In their overview of Berber languages, Aikhenvald & Militarev (1991:218) mention the languages with a single ‘who?, what?’ interrogative from the Northern, Western and Eastern branches. For instance, they report that the interrogative *ma* ‘who?, what?’ can be found in the so-called “continuous” Zenati languages (including the East Zenati subgroup) and Ntifa (Tachelhit group) from the Northern branch. These languages are spoken primarily in Morocco, Algeria and Tunisia. In Figure 3, I list (most of) the idioms classified by Aikhenvald & Militarev (1991:158-159) as “continuous” Zenati. I also give the Atlas Berber languages (Tachelhit and Tamazight groups) here because some of them will be mentioned later in the discussion. The idioms classified as Zenati of “oases” are not presented in full because they are less relevant for the discussion below.

²⁸ I would like to thank Lionel Galand, Karl-G. Prasse and Catherine Taine-Cheikh for sharing their data with me and for discussion. I am also grateful to Maarten Kossmann for his comments on the present section.

Figure 2. Classification of modern Berber languages (based on Aikhenvald & Militarev 1991:157-159)

Eastern Berber (genetic unity is questionable)

- Siwa (Siwa oasis, Egypt)
- Awjila (Jalu oasis, Libya)
- Ghadamès (Ghadamès oasis, Libya)

Southern Berber (or **Tuareg**)

- North Tuareg (Algeria, Libya, Mali and Niger, also in Nigeria and Chad)
- South Tuareg (Mali and Niger, also in Burkina-Faso and Nigeria)

Western Berber (or **Zenaga**; south-west Mauritania)

Northern Berber

- Atlas Berber
 - Tachelhit (or Shilha, Chleuh) group (Morocco, Algeria)
 - Tamazight (or Beraber) group (Morocco, Algeria)
- Zenati Berber
 - “Continuous” Zenati (Morocco, Algeria, Tunisia)
 - Zenati “of oases” (Morocco, Algeria, Libya)
- Kabyle Berber

Note that in all probability, not all “continuous” Zenati idioms presented in Figure 3 have a ‘who?, what?’ interrogative. For instance, Tamezret, spoken in Southeast Tunisia, does not have such an interrogative (cf. Ben Mamou 2005a). I have not been able to consistently check all the idioms.

Aikhenvald & Militarev report Zenaga (Western Berber; Mauritania) to use the form *min* as a ‘who?, what?’ interrogative. I have not been able to establish the variety of Zenaga Aikhenvald & Militarev (1991) had in mind because the interrogative of the form *min* is lacking in the sources on Zenaga I consulted (Basset 1909, Faidherbe 1877, Nicolas 1953; Catherine Taine-Cheikh, p.c.).²⁹ Aikhenvald & Militarev also mention the Berber language of Ghadamès oasis in Libya (Eastern Berber according to their classification, Northern Berber of the East Zenati group according to the Ethnologue) to have an interrogative *inu* ‘who?, what?’. However, this is not supported by the sources I consulted, de Calassanti-Motyliniski (1904) and Lanfry (1968, 1973), which are also the only sources on Ghadamès mentioned in Aikhenvald & Militarev’s (1991) references.

²⁹ In Nicolas (1953:205) there is only *mi:n* ‘man, person’, *mā:n* ~ *man* ‘people’. There remains only one source on Zenaga in Aikhenvald & Militarev’s (1991) references that I did not manage to consult, Masqueray (1879).

Figure 3. Northern Berber groups with languages having a 'who?, what?' interrogative (based on Aikhenvald & Militarev 1991:158-159, 218)³⁰

Zenati Berber

“Continuous” Zenati (Morocco, Algeria, Tunisia)

Ait Seghrouchen

Ghomara,ⁱ⁾ Jbalaⁱ⁾

Rif/Tarifit

Senhajaⁱ⁾

Ait Ouarain

Beni Iznasen

Tlemcen region varieties: Beni Snous, Beni bou Said

Oued Cheliff basin varieties: Matmata, Haraoua, Achacha, Halima, Beni bou Attab, Beni

Ferah, Senfita, Gheraba; Beni Menacer; Chenoua; Beni Salah,ⁱⁱ⁾ Beni Messaoud,ⁱⁱ⁾

Beni Misraⁱⁱ⁾

Tachawit/Chaouia

East Zenati: Djerba island varieties (Ajim, Guellala, Sedouikech, Al Ma'i, Mahboubin,

Sedghiane); Southeast Tunisia varieties (Zrawa, Taoujjout, Tamezret, Chenini,

Douiret); Zuara

Zenati of “oases” (Morocco, Algeria)

Mzab

Ouargla

Figuig

Oued Righ varieties

Gourara

etc.

Atlas Berber

Tachelhit (or Shilha, Chleuh) group (Morocco, Algeria)

Drâa River basin & Jebel Bani mountain region varieties: Tinduft, Ait Umribed, etc.

Anti-Atlas varieties: Izemdaln, Imitek, Imejjad, Illalene, Ida ou Kensous, Ida ou Zekri,

Ida ou Zeddout, Amanouz, Akhsas, Ait Mzal, Ait Isaffen, Ida Gnidif, Tagountaft,

Iglioua, Ait Ouazouguit, etc.

Sous River basin varieties: Tazeroualt, Ait Baamran, Haouwara, Ida ou Semlal,

Achtouken, Aksimen, Masst, Talekjount, Tigouga, Seksaoua, Mentaga, Ait Ouajes,

Ida ou Mhamoud, Ida ou Izimmer, Demsira, etc.

High Atlas varieties: Touggana, Iguedmioun, Ait Immour, Ihahan, Imeghran, Ida ou

Tanan, Ida ou Zikki, Ida ou Kais, Ida ou Zal, Intift/Ntifa, etc.

³⁰ The names of the Berber languages in Figure 3 are transliterated from Russian, but I tried to use one of the common French or English writings whenever I managed to identify one. Often, the names of the varieties are just the names of the locations where they are spoken.

Tamazight (or Beraber) group (Morocco, Algeria)

Demnat region varieties: Ait Messad, etc.

Varieties of the upper reaches of the Dadès River and High Atlas: Ait Izdeg, Ait Yahya, Ait Sliman, Ait Khebbach, Ait Yafelman, etc.

Middle Atlas varieties: Ait Sadden, Ait Yusi, Izayan, Ait Sgougou, Ichkern, Ait Ishak, Ait Mgild, etc.

Meknès region varieties: Ait Ndir, Ait Naaman, etc.

Notes:

i) According to Maarten Kossmann (p.c.), these languages are not Zenati.

ii) According to Maarten Kossmann (p.c.), these languages should rather be classified as Kabyle.

An interrogative of the form *inu* is given only by de Calassanti-Motyliniski (1904:27), where it is glossed as ‘who?’, while Lanfry (1968:357, 1973:230) gives only *anno* ‘who?’.

I found the following Berber idioms with a ‘who?, what?’ interrogative. In Ait Seghrouchen of Oum Jeniba this interrogative has the forms *may*, *magg*³¹ and *ma*, depending on the morphosyntactic context (Bentolila 1981:101-111). Ait Seghrouchen is usually considered to belong to the Tamazight group of the Northern branch, but Aikhenvald & Militarev (1991:156) rather believe it to be a “continuous” Zenati language. In Ait Ndir, a Northern Berber language of the Tamazight group, the ‘who?, what?’ interrogative has the forms *may*, *mag(-g)*, *ma*, *m*, *mi* and probably also *matta* (Penchoen 1973:79-82; cf. also Section III.2.3.2.2.5). The Sous River basin variety of Tachelhit (Northern Berber; Morocco) described by Applegate (1958:28) may be using the form *ma* as ‘who?, what?’, but the information on interrogative pronominals is somewhat incoherent. In Basset’s (1909:25) description of a variety of Zenaga, the form *ma* is glossed as ‘who?, what?’. Catherine Taine-Cheikh (p.c.) reports the interrogative *ta’k(k)* ‘who?, what?, which one?’ for the Idab Lahsan variety of Zenaga. Hanoteau (1896:47-49) mentions the form *ma* ‘who?, what?’ for a variety of Ajjer or Ahaggar Tuareg (Southern Berber, North Tuareg; Algeria). De Calassanti-Motyliniski & Basset (1908:317-318) also give *ma* ‘who?, what?’ for an Algerian variety of Tuareg, which I have not been able to identify.

Heath (2005) reports another interesting overlap in the domain of interrogative proforms for Malian varieties of Tuareg. One and the same form *mi* is used there as ‘who?’ and as ‘where?, whence?’ (next to the specialized ‘where?’ interrogatives *əndek*, *əndəke* and *ənde*), as illustrated in (43) and (44).

³¹ What some sources treat as geminates, others consider to be so-called “tense” consonants. These “tense” consonants are usually represented with capitalized consonantal letters. For the sake of uniformity, I have opted for the geminate notation.

Malian Tuareg (Southern Berber, South Tuareg; Mali)³²

- (43) *mí-dd Ø-ðsæ-n?*
 IPW-VEN 3SG.M.SBJ-arrive.PFV.POSITIVE-PTCP.M.SG
 ‘Who came?’ (Heath 2005:650)
- (44) *mí-dd i-ffál?*
 IPW-VEN 3SG.M.SBJ-leave.RES
 ‘Where does he come from?’ (Heath 2005:653)

However, in the meaning ‘where?’, it is used “only with transitive motion verbs that take a complement denoting either the terminus or the point of origin” (Heath 2005:653). The danger of confusion between the two meanings is further reduced due to the necessary use of a participle form of the verb in (43), which clearly indicates that *mí* is the subject (cf. Section III.2.3.2.2) and, consequently, means ‘who?’, whereas the non-participle form of the verb in (44) indicates that *mí* is the complement of a transitive motion verb and, consequently, means ‘where?, whence?’.

2.3.2.2 Berber interrogative pronominals as constructions: simplicity of structure & multitude of forms

2.3.2.2.1 Preliminaries

A description of a Berber language would often reveal a rather large number of forms of interrogative pronominals. In older descriptions, the reader is usually just confronted with a set of forms without any additional explanations, as, for instance, in Basset’s (1909) sketch of Zenaga where the interrogative pronominals cited in (45) are given without any comment.

Zenaga (Western Berber; Mauritania & ?Senegal; Basset 1909:25)

- (45) ‘what?’ *ta, kai, keid, id’i*
 ‘who?’ *taka*
 ‘who?, what?’ *ma*
 ‘which?, what (kind of)?’ *matak*

However, more recent descriptions make it clear that the large inventories of Berber interrogative pronominals are primarily due to the fact that typically, interrogative pronominals in Berber languages are not single morphemes, but

³² Accentuation marks in Malian Tuareg examples do not represent phonetic differences. Instead, they distinguish a lexically determined stress, marked with an acute accent, from the stress determined by various “default accentuation” rules, which is marked with a single or double grave accent (Heath 2005:20-21). I do not use accents in forms cited in isolation to prevent possible misinterpretations of the degree of their relevance for lexical differentiation.

constructions. The construction at issue is similar to the one used for the purposes of focalization, often referred to as “(complex, pronominal or reinforced) anticipation” in Berber studies. In turn, focus constructions are formally very similar and clearly related to relative constructions, although in modern languages they tend to become differentiated in some details (cf. Aikhenvald & Militarev 1991:237 and Heath 2005:623-648, among others). The focalized element is fronted to sentence-initial position and followed by a relative clause, which is introduced by a demonstrative pronominal. This construction can be analyzed as a kind of cleft. When the extracted element is subject, the verb of the relative clause is in participle form and often takes the invariable third person singular masculine subject marker. That is, it does not agree in gender and number with the extracted subject.

Demonstrative pronominals, which constitute an important part of the construction at issue, form very rich paradigms in Berber languages (similarly to many other Afro-Asiatic languages), distinguishing among other things various degrees of distance, grammatical gender, number, anaphoric reference, humanness and definiteness (cf. Basset 1952:34-35). Demonstrative modifiers form smaller paradigms, with generally three members distinguishing only distance and anaphoric reference. Usually, there are also demonstrative pronominal roots unmarked for gender-number, such as *ā* in Malian Tuareg (Heath 2005:239) and *ay* in Ait Seghrouchen (Bentolila 1981:96) and Ait Ndir (Penchoen 1973:23). Thus, in Ait Ndir, the demonstrative pronominal root *ay* is described by Penchoen (1973:23) as referring to “actions, things said, things of which gender is unknown, etc.”. If agreement of the verb is necessary, the verb will be in the third person singular masculine, which is thus the default option. The demonstrative pronominal root *ay* can also combine with demonstrative modifiers producing such gender-number neutral forms as *ay-ad* ‘this’, *ay-inn* ‘that’ and *ay-nna* ‘that, in question’. Typically for Northern Berber languages of the Kabyle and Tamazight groups, demonstrative modifiers are gender-number insensitive (Aikhenvald 2000:39). Malian Tuareg frequently uses gender-number insensitive demonstrative modifiers as well (Heath 2005:241). Furthermore, in Malian Tuareg, the gender-number neutral demonstrative pronominal *a* can occur preceding the masculine singular proximal, distant and unmarked demonstrative pronominals to “form a neutral demonstrative glossable ‘this’, ‘that’, with no specific gender identity: *a wa* ‘DEM’ (unmarked), *a wahi* ‘this’, *a wen* ‘that’” (Heath 2005:239). In Berber languages, demonstrative pronominals unmarked for gender-number are especially common means of introducing the relative clause in focalized constructions and consequently, in constituent questions. They can also be used to introduce ordinary relative clauses, although in the latter case there is also a tendency for some regular gender-number marked

demonstrative pronominals to become generalized irrespectively of the gender-number of the antecedent (cf. Aikhenvald & Militarev 1991:236).

Summing up, (i) interrogative pronominals in Berber languages often appear to originate in cleft constructions composed of an interrogative pronominal and a demonstrative pronominal, (ii) gender-number distinctions in the latter demonstrative may be important for the differentiation between the features person and thing of the whole construction, but (iii) demonstratives introducing the relative clause in focalized constructions and consequently, in constituent questions are often gender-number neutral. In what follows, I will examine the structure of various Berber interrogative pronominals. For convenience sake, interrogatives are grouped together in sections according to formal similarities.

2.3.2.2.2 The interrogatives *ma*, *may*, *mad*, and the like

‘Who?, what?’ interrogatives such as Ait Seghrouchen and Ait Ndir *may*, *mag(g)* and *ma* consist of a single interrogative root *m* ‘who?, what?, which one?’ and the forms of the gender-number neutral demonstrative pronominal *ay/agg/a*, as in (46-48) from Ait Ndir.

Ait Ndir (Northern Berber, Atlas Berber, Tamazight; Morocco; Penchoen 1973:79)

- (46) *m-ag-gæddan* *ɣər-taddart?*
 IPW-DEM-3SG.M.SBJ.go.PFV.PTCP.SG to-house
 ‘Who has been to the house?’
- (47) *m-ay nttannay asəkka?*
 IPW-DEM FUT.1PL.SBJ.see tomorrow
 ‘Whom/what shall we see tomorrow?’
- (48) *m-a mi iša lflus?*
 IPW-DEM to 3SG.M.SBJ.give.PFV money
 ‘To whom did he give the money?’

Demonstrative pronominals may have different forms conditioned by their morphosyntactic environment (or syntactic function). For instance, in Ait Ndir, the gender-number neutral demonstrative pronominal *ay* has the form *a* in combination with adpositions and it fuses in the form *agg* with the following third person singular masculine verbal subject prefix *i-*, as illustrated in (49-51).

Ait Ndir (Penchoen 1973:78)

- (49) *irəyzən ay nannay asa-ənnat*
 men DEM 1PL.SBJ.see.PFV yesterday
 ‘It’s the men who we saw yesterday’

- (50) *irəyzən a mi iša lflus*
 men DEM to 3SG.M.SBJ.give.PFV money
 ‘It’s the men to whom he gave the money’
- (51) *asa-ənnat aggda γər-ssuq*
 yesterday DEM.3SG.M.SBJ.go.PFV to-market
 ‘It was yesterday that he went to market’

In Ait Seghrouchen, a similar gender-number neutral demonstrative pronominal *ay* has the form *a* when followed by the irrealis marker *γa*, negation marker *ur* and certain adpositions and it fuses in the form *agg* with the following third person singular masculine verbal subject prefix *i-/y-* (Bentolila 1981:108).

In Ait Ndir, the interrogative root *m* can be directly juxtaposed to a nominal in just a few cases, as in (52) with *ism-ənnəs* ‘his/her/its name’.

Ait Ndir (Penchoen 1973:79)

- (52) *m ism-ənnəs?*
 IPW name-GEN.3SG
 ‘What’s his/her name? What’s it called?’

Another example is provided by *m-imš* ‘how?’, which according to Penchoen (1973:80), is clearly related to *imši-s* ‘thus, this way’.³³

According to Lionel Galand (p.c.), in varieties of Tachelhit, the gender-number neutral demonstrative pronominal *ad* has the form *a* when followed by the third person singular masculine verbal subject prefix *i-*, the first person plural verbal subject prefix *n-* or the allative particle *nn*, it has the form *as* when followed by the third person singular feminine encliticized object pronoun *tt*, and it fuses into the forms *add*, *akk*, *att* and *ayy* with the morphemes beginning with *d-*, *k-*, *t-* and *y-*, respectively. Consequently, the interrogative ‘who?, what?’ has the forms *mad*, *madd-*, *makk-*, *matt-*, *mayy-*, *mas* and *ma* in this language, (53-54).

Tachelhit (Northern Berber, Atlas Berber; Morocco; Lionel Galand, p.c.)

- (53) *m-att-zrit?*
 IPW-DEM.2SG.SBJ-see
 ‘Who/what do you see?’

³³ Penchoen (1973:80-81) mentions a few other interrogatives beginning with *m-*, such as *milmi* ‘when?’, *mšta* ‘how much?, how many?’, *mahallig* ‘why?’ and *mani* ‘where?’, which he believes to be compounds etymologically, but for which he is unable to indicate the noun that served as the second part of the compound. The interrogative *mani* ‘where?’ will be discussed in more detail in Section III.2.3.2.2.4.

- (54) *m-a igan amyar?*
 IPW-DEM 3SG.M.SBJ.be.PTCP.SG chief(M)
 ‘Who is the chief?’

In Zenaga of Idab Lahsan, according to Catherine Taine-Cheikh (p.c.), there is a specifically human interrogative *m-ād* ‘who?’ consisting of the interrogative root *m*, found only in this collocation, and the proximal human masculine singular demonstrative pronominal *ād* ‘this’, as in (55).

Zenaga of Idab Lahsan (Western Berber; Mauritania)

- (55) *m-ād y-ūdiž-ān?*
 IPW-DEM.M.SG.HUM 3SG.M.SBJ.lie.down.PFV-PTCP.SG
 ‘Who lay down?/ Who is lying? (the original French translation: ‘Qui (s’)est couché?’)’ (Catherine Taine-Cheikh, p.c.)

Ahaggar Tuareg, as described by de Foucauld (1951-1952), has two interrogative pronominals *m-i* ‘who?’ (56a) and *m-a* ‘what?’ (56a) or ‘who?’.³⁴

Ahaggar Tuareg (Southern Berber, North Tuareg; Algeria; de Foucauld 1951-1952 under *mi* and *ma*)

- (56) a. *m-i tenēyed řur-es?*
 IPW-DEM 2SG.saw by-3SG
 ‘Who did you see at his place?’
 b. *m-a tenēyed řur-es?*
 IPW-DEM 2SG.saw by-3SG
 ‘What did you see at his place?’

Note that *m-a* can mean ‘who?’ (almost) only in combination with an indicative present form of the verb *umas* ‘be’,³⁵ as in (57a), while the same question with *m-i* is a KIND-question about the person’s affiliation as regards his or her family, tribe, nationality, etc., as in (57b).

³⁴ I present the two Ahaggar Tuareg interrogatives as bimorphemic by analogy to the identical Air Tuareg interrogatives *m-i* and *m-a* (cf. Section III.2.3.2.2.3).

³⁵ I say *almost* because de Foucauld (1951-1952, under *ma*) remarks that sometimes *ma* can be heard to be used instead of *mi*, although according to him, it is an impropriety (“c’est une incorrection”). Furthermore, as was mentioned in Section III.2.3.2.1, Hanoteau (1896:47-49) and De Calassanti-Motylinski & Basset (1908:317-318) also report *ma* ‘who?, what?’ for a variety of Ajjer or Ahaggar Tuareg and an unidentified Algerian variety of Tuareg respectively and they do not mention any restrictions on the use of *ma* as ‘what?’ of the kind described by de Foucauld (1951-1952).

Ahaggar Tuareg (de Foucauld 1951-1952 under *mi* and *ma*)

- (57) a. [A:] *m-a* *temûsed?* [B:] *nek Kenân*
 IPW-DEM 2SG.be.PRS.IND 1SG PROP
 ‘[A:] Who are you? [B:] I am Kenân (*Kenân* is a personal proper name)’
- b. [A:] *m-i* *temûsed?* [B:] *Kel-Ṛela*
 IPW-DEM 2SG.be.PRS.IND PROP
 ‘[A:] What are you? (Of what people are you?) [B:] I am a Kel-Ṛela (a member of the Kel-Ṛela tribe)’

It seems that *m-a* with *umas* can sometimes be used as a KIND-question as well. Although de Foucauld (1951-1952) does not provide any examples of such use, this possibility follows from his description.³⁶

The use of *m-a* with *umas* ‘be’ as a non-human ‘what?’ is illustrated in (58).

Ahaggar Tuareg (de Foucauld 1951-1952 under *ma*)

- (58) a. *m-a* *yemûs* *awah?*
 IPW-DEM 3SG.M.be.PRS.IND DEM
 ‘What’s this?’
- b. *m-a* *yemûs* *ăbarekka wa tegïd?*
 IPW-DEM 3SG.M.be.PRS.IND way DEM 2SG.did
 ‘What is the way that you took?’

2.3.2.2.3 The interrogatives *mi* and *wi*

The interrogative pronominal of the form *mi* usually means ‘who?’ in Berber languages, as, for instance, in Air Tuareg (North Tuareg; Niger; Galand 2002 [1974]) or Ahaggar Tuareg (North Tuareg; Algeria; Karl-G. Prasse, p.c.). However, it can also mean ‘who?, where?’, as in Malian Tuareg (South Tuareg; Mali), ‘who?, what?’, as in Ait Ndir (Northern Berber; Morocco), and maybe ‘what?’, if the Ghadamès (Eastern or Northern Berber; Libya) interrogative *me* (Lanfry 1973:191) also belongs here. The usage of the interrogative pronominal *mi* in Ait Ndir is illustrated in (59).

³⁶ Thus, he writes “*ma*, suivi de l’ind[icatif] prés[ent] du v[erbe] *umas* ‘être’, signifiant ‘que suis-je? qu’es-tu? qu’est-il? etc. [...] a souvent le s[ens] ‘qui suis-je? qui es-tu? qui est-il? [...]’ [emphasis added]”. The use of *souvent* ‘often’ implies here that sometimes, *ma* with *umas* still can mean ‘what am I?’, ‘what are you?’, etc.

Ait Ndir (Penchoen 1973:82)

- (59) a. *taddart-ad t-i-n-mi?*
 house.F-this F-DEM-GEN-IPW
 ‘Whose house is this? (lit.: ‘This house is the one of who’)
- b. *s-mi?*
 with-IPW
 ‘with what/whom?’
- c. *xf-mi?*
 on-IPW
 ‘on what/whom?, about whom/what?’
- d. *yer-mi?*
 to-IPW
 ‘to(ward) who? (to whose place?)’
- e. *šəgg d-mi?*
 2SG.M with-IPW
 ‘you and who?’

Note that *mi* in Ait Ndir is not used on its own and the preference for one or another reading is to a considerable extent determined by the semantics of other elements of the construction it makes part of, such as the preposition, the verb, etc. In this respect, the *mi* interrogative of Malian Tuareg, illustrated in (43-44), does not differ that much from the *mi* interrogative of Ait Ndir. For instance, *mi* in Malian Tuareg example (44) could have also been glossed as ‘what?’, so that its ‘where?’ reading is just due to the fact that it is used with a transitive motion verb that takes a “complement denoting [...] the point of origin”.

The interrogative pronominal *mi* can also be analyzed as consisting of two parts, an interrogative *m* and a demonstrative pronominal root *i*.³⁷ This is, for instance, what Galand (2002:123-124 [1974]) does for Air Tuareg, where interrogatives *m-i* ‘who?’ (60a) and *m-a* ‘what?’ (60b) can be found.

Air Tuareg (Southern Berber, North Tuareg; Niger; Galand 2002:124 [1974])

- (60) a. *m-a dak-əge?*
 IPW-DEM to.you-I.do
 ‘What do I do to you?’

³⁷ Note that *-i* in the Malian Tuareg form *mi* ‘where?’ may also represent the common Berber locative case marker *-i* or adposition *i* ‘to’ (cf. Aikhenvald & Militarev 1991:212 & Section III.2.3.2.2.4). However, taking into consideration the usual structure of Berber interrogative pro-words, I am reluctant to accept this hypothesis.

- b. *m-i* *ilən* *tatte?*
 IPW-DEM PTCP.M.SG.have.PTCP.M.SG food.F
 ‘Whose is the food?’

The element *a* is apparently cognate to the gender-number neutral demonstrative pronominals of similar shapes illustrated above for Ait Ndir, Ait Seghrouchen, Tachelhit and Malian Tuareg (see Section III.2.3.2.2.1-III.2.3.2.2.2). As in other languages, in Air Tuareg this gender-number neutral demonstrative pronominal root is also used to build focus constructions, (61).

Air Tuareg (Galand 2002:124 [1974])

- (61) *ehəre-nin* *a* *yæggen*
 herd(M)-GEN.1SGDEM PTCP.M.SG.be.numerous.PTCP.M.SG
 ‘It is my herd that is (the most) numerous’

Besides the combination with the interrogative *m*, the element *i* is found in Air Tuareg as “indefinite” (non-specific, free-choice) masculine (singular or plural) relative demonstrative pronominal, with *t-i* as its feminine counterpart, (62a, b).

Air Tuareg (Galand 2002:123 [1974])

- (62) a. *ehəre-nin* *i* *yæggen*
 herd(M)-GEN.1SG DEM PTCP.M.SG.be.numerous.PTCP.M.SG
 ‘My herd is numerous (lit.: ‘My herd is one that is numerous’)
- b. *tatte* *t-i* *n* *kullu-nnəwən*
 food(F) F-DEM GEN each-GEN.2PL
 ‘The food is of you all (lit.: ‘The food is one of each of yours’)

Interestingly, in Malian Tuareg similar “indefinite” relative demonstrative pronominals *i* and *t-i* are restricted to human referents, as in (63a, b), and *a* to non-human referents, as in (64a) (except when the latter is used to create a focus construction, as in (64b)).

Malian Tuareg

- (63) a. *i-dád* *osæ-nen*
 DEM-VEN go.RES-PTCP.PL
 ‘those(M) who (= whoever) have come’ (Heath 2005:639)
- b. *t-i-dád* *t-osæ-t*
 F-DEM-VEN PTCP.F.SG-go.PFV.POSITIVE-PTCP.F.SG
 ‘that one (F) who (= whoever) comes’ (Heath 2005:639)
- (64) a. *à* *t-əjræw-æd*
 DEM 2SBJ-find.RES-2SG.SBJ
 ‘what(-ever) you (M.SG) find (lit. ‘have found’)

- b. *nəkk à i-ssán-æn*
 1SG DEM PTCP.M.SG-know.RES-PTCP.M.SG
 ‘It is I who knows’ (Heath 2005:644)

Malian Tuareg also uses *i* when the relative clause modifies a pronoun of the first or second person, (65).³⁸

Malian Tuareg (Heath 2005:624)

- (65) *nəkk-æn-æt-eḍ i əkšə-nen*
 1PL.F DEM eat.RES-PTCP.PL
 ‘we (F) who have (already) eaten’

Furthermore, the marker *i* is widely used in Berber languages to mark plurality of gender-number sensitive demonstrative pronominals, as in Malian Tuareg (Heath 2005:239) and Ait Ndir (Penchoen 1973:22). In certain syntactic contexts, it is also not uncommon for this marker to be used for purposes of number neutralization in gender-number sensitive demonstrative pronominals, as in Ait Ndir examples (66) and (59a).

Ait Ndir (Penchoen 1973:23)

- (66) a. *w-i-nnəs*
 M-DEM-GEN.3SG
 ‘his/her masculine-one(s)’
 b. *t-i-nsən*
 F-DEM-GEN.3PL.M
 ‘their (M) feminine-one(s)’
 c. *w-i-s-sin*
 M-DEM-with-two.M
 ‘the second masculine-one(s)’
 d. *t-i-s-snat*
 F-DEM-with-two.F
 ‘the second feminine-one(s)’

³⁸ Heath (2005) treats the demonstrative pronominal *i* used with first and second person pronouns as different from the indefinite *i* because it does not have a feminine form and is necessarily definite. I do not question the soundness of this analysis. However, I would like to point out that there are similarities between the two *i* demonstrative pronominals. In particular, what the inherently definite and indefinite uses share is that in both cases the referents are “not in need of referential disambiguation”, either because they are already inherently definite or because they cannot be disambiguated (the terminology is due to Van de Velde 2006, see Section III.2.1.1.2.1 for more details).

It is not implausible that the pluralizing and number neutralizing uses are also related to the “indefinite” (non-specific, free-choice) uses of *i*, such as those already mentioned for Malian Tuareg or as in the case of Chaouia (Northern Berber; Algeria) demonstrative pronominals *w-a* ‘exactly this (M.SG)’ vs. *w-i* ‘any one of these (M.SG)’ (Aikhenvald & Militarev 1991:217). Otherwise, the “indefinite” *i* could be cognate to the somewhat similar looking numeral root ‘one’, which would then have been incorporated into the system of demonstrative pronominals with subsequent formal and structural adjustments to its new paradigm.³⁹ The specialization of the “indefinite” *i* for human reference in Malian Tuareg mentioned above may then be compared to the exclusively human reference of *someone* in English. For another possible source of *-i*, namely case marking see Section III.2.3.1.

Interestingly, in various Northern Berber languages the “indefinite” masculine demonstrative pronominal *wi*, such as the Chaouia *w-i* ‘any one of these (M.SG)’ mentioned above, appears to develop into a human interrogative pronominal ‘who?’, e.g. Kabyle Berber, as in (67), Figuig, (68), and Ait Seghrouchen.

Kabyle Berber (Northern Berber; Algeria; Lionel Galand, p.c.)

- (67) *wi* *ibyan* *aman*
 INDF.M.SG/IPW PTCP.want.PFV.PTCP water
 ‘Who(ever)/ the one who wants some water...’ or with interrogative intonation ‘Who wants some water?’

Figuig (Northern Berber, Zenati “of oases”; Morocco; Kossmann 1997:201)

- (68) a. *wi* *xef-sent* *i-jawb-en* *ad* *y-away*
 INDF.M.SG on-3SG.F PTCP-answer.PFV-PTCP FUT 3SG.M.SBJ-bring.AOR
yelli-s
 daughter-GEN.3SG
 ‘The one who (whoever) will answer, will marry his daughter’
 b. *wi* *ked* *t-urər-ed*
 IPW with 2SG.SBJ-play.PFV-2SG.SBJ
 ‘With whom did you play?’

³⁹ Here are the forms of the numeral ‘one’ in some Berber languages, with the masculine form preceding the feminine one: Kabyle (Northern Berber) *iin* & *iat* (Aikhenvald & Militarev 1991:214), *yiwen* & *yiwat* (Basset 1952:28), Tachelhit (Northern Berber) *yan* & *yat* (Basset 1952:28), Tamazight (Northern Berber) *idž* & *ist* (Aikhenvald & Militarev 1991:214), Chaouia of Ait Frah (Northern Berber) *yəžž(ən)* & *tišt* (Basset 1952:28-29), Siwa (Eastern Berber) *əgən* & *əgt* (Aikhenvald & Militarev 1991:214), Malian Tuareg (Southern Berber) *iyæn* & *iyæt* (Heath 2005:248), Zenaga (Western Berber) *iun* & *nəiiun* (Aikhenvald & Militarev 1991:214).

The link between the demonstrative pronominal and the interrogative is also supported by the fact that in some languages, such as Tamezret, both elements appear to have become fixed with the suffixed distant demonstrative modifier *-(i)n ~ -(e)n*.⁴⁰ Thus, in Tamezret both elements have the form *win*, (69) which can be compared to *win* ‘that one (M.SG)’ of Ait Seghrouchen (Bentolila 1981:53, 93) or *wen* (<**w-a-en*) ‘that one (M.SG)’ and *win* (<**w-i-en*) ‘those ones (M.PL)’ of Malian Tuareg (Heath 2005:239-240).

Tamezret (Northern Berber, “continuous” Zenati, East Zenati; Tunisia)

- (69) a. *d win i yəswə*
 it.is IPW DEM 3SG.M.SBJ.drink.PST
 ‘Who drank?’ (Ben Mamou 2005b:8)
- b. *win ədyəs d əməzwər əyxalləs*
 INDF.M.SG FUT.3SG.M.SBJ.arrive it.is first.M FUT.3SG.M.SBJ.pay
 ‘Whoever will arrive, it is the first who will pay (or ‘The one who arrives first will pay’)’ (Ben Mamou 2005b:12)

Furthermore, a somewhat similar development from a masculine demonstrative pronominal to an interrogative pronominal has been claimed for Middle Egyptian *pw* (cf. Section III.2.3.3.2.2). Frajzyngier (1985) reconstructs for Proto-Chadic an interrogative **wa* ‘who?’, which suspiciously resembles the widespread Berber masculine demonstrative pronominal *wa*. All these coincidences taken together seem unlikely to be due to chance only.

In all probability, the Ghadamès interrogative *inu* ‘who?’ (de Calassanti-Motyliniski 1904:27) also belongs here. It may be compared to the Figuig (Kossmann 1997:193) and Ait Seghrouchen (Bentolila 1981:93) masculine plural proximal demonstrative pronominals *inu* and *yinu*, respectively.

2.3.2.2.4 The interrogatives *man*, *mani*, *an*, and the like

The interrogative proform of the form *man* usually means ‘which [N]?, what (kind of) [N]?’ in Berber languages, as in Ait Seghrouchen *man aryaz* ‘which man?’ (Bentolila 1981:33), in Ahaggar Tuareg *man-əkət* ‘how much?, how many?, how long (time)? (lit. ‘what measure’)’ (Karl-G. Prasse, p.c.) or in Tachelhit *man-tifii* ‘what (kind of) meat?’ (Applegate 1958:28). In Senhaja, *man* has acquired the meaning ‘who?’ (Aikhenvald & Militarev 1991:218), the development from ‘which one?’ to ‘who?’ being quite natural (cf. Section I.2.3).

⁴⁰ It may be more correct to speak of a “non-proximal” demonstrative modifier here rather than of a simply “distant” one, because various degrees of distance can sometimes be distinguished in Berber languages.

Like the interrogatives *ma* and *mi*, the interrogative *man* can also be analyzed as consisting of at least two parts, *m* and *an*. This is suggested by the comparison of such forms as Tachelhit *man wa* ‘which one.M? (lit. ‘which this.M.SG’)’ (Lionel Galand, p.c.) and Kabyle *an wa* ‘which one.M? (lit. ‘which this.M.SG’)’ (Hamouma 1987:70). In all probability, the Ghadamès interrogative *anno* ‘who?’ (Lanfry 1973:230) can be considered here as well. For Malian Tuareg Heath (2005:660-661) compares the interrogative *man-iket* ‘how much?, how many?’ with its counterpart in the relative clause head function *an-iket* ‘(I don’t know) the amount (he paid)’ and comes to the conclusion that *man*-historically contains an interrogative *m*. Further evidence comes from a comparison with a widespread interrogative *mani* ‘where?’, which is clearly related and is quite commonly used in the meaning ‘which (one)?’ as well, e.g. Ait Seghrouchen *mani tawriqt* ‘which leaf?’ (Bentolila 1981:196) or Niger varieties of Tuareg *mani-tu* ‘which is he?, where is he?’ (Karl-G. Prasse, p.c.). According to Aikhenvald & Militarev (1991:212), the final *-i* in *mani* is the common Berber locative case marker (it may also be cognate to the widespread Berber adposition *i* ‘to’). The initial *m-* in *mani* can be singled out rather easily as well. Consider, for instance, the following interrogative proadverbs: Kabyle (Hamouma 1987:70) *anida* ~ *anda* ‘where?’ (with *da* ‘here’) and *sani* ‘where from?’ (with the adposition *s* ‘from’), and Malian Tuareg (Heath 2005:653) *andek* ‘where?, which (one)?’. Penchoen (1973:80) also speaks about “a number of Berber dialects [that] have a word *ani* ‘where, the place in which’”.⁴¹

As various other interrogatives considered in Sections III.2.3.2.2.2-III.2.3.2.2.3, the *an* part of (*m*)*an* may represent a demonstrative pronominal. A possible cognate could be, for instance, the Zenaga non-human distal demonstrative pronominal *ān* (Catherine Taine-Cheikh, p.c.) or the Egyptian (subject focus) particle *jn* */(ʔ)an(V)/ ‘it is [N]’ (cf. Allen 2000:189, Gardiner 2001:175-177 [1957], Loprieno 1996:134-136), which according to Gardiner (2001:53 [1957]) is “probably demonstrative in origin”.⁴² However, it might also

⁴¹ It should be mentioned that in certain cases ‘where?’ interrogatives containing the part *an* are likely to be of a different origin. In particular, I have in mind such forms as Kabyle *ansi* ‘where to? (with *i* ‘to’)’ (Hamouma 1987:70) and, probably, Tachelhit of Tazeroualt *manza* ‘where?’ (Stumme 1899:120 via Aikhenvald 1995a:63). These forms seem to be based on the noun meaning ‘place’, which, for instance, in Ait Seghrouchen has the form *ansa* and can be used in combination with *man* ‘which?’ as *man ansa yr* ‘where to? (with *yr* ‘to’)’ (Bentolila 1981:26).

⁴² That the Egyptian particle *jn* should be read as something like */(ʔ)an(V)/ is suggested by its Coptic reflexes *an* ~ *n*, for instance, in the so-called “independent” (or “stressed”) pronouns (cf. Loprieno 1996:108-109). When used word-initially, the sign *j* (or *ʔ*) in Egyptian transliterations stands either for a vowel, perhaps preceded by a glottal stop /ʔ/, or for a palatal approximant /j/ (Gardiner 2001:28-29 [1957], Callender 1975:3). Lipiński (1997:476-477) mentions the Egyptian *jn* as a cognate of a Proto-Semitic pronominal basis *²*an* ~ *na*. Compare also various

be possible to split *(m)an* even further by singling out the final *n*, which would then be the common Berber genitive (or possessive) construction marker *n*, the so-called *nota genitivi*, also often described as a preposition ‘of’. Thus, de Calassanti-Motyliniski & Basset (1908:317-318) analyze the Algerian Tuareg interrogative *man-eket* ‘how much?, how many?’ (with *eket* ‘measure’) as *ma-neket* ‘what of the measure’. This would account for the typically attributive use of *(m)an*. However, the above analysis is somewhat complicated by the fact the genitive *n* ‘of’ regularly requires the dependent noun (i.e., the possessor) to be in the so-called “annexed (or construct) state” (*état d’annexion* or *état construit*), whereas the noun following *(m)an* normally seems to occur in the so-called “free state” (*état libre*).⁴³ For instance, in Ait Seghrouchen one finds *man aryaz* ‘which man?, what man?’ (Bentolila 1981:33), where the free form *aryaz* and not the annexed form *uryaz* is used.

It is possible to get around this problem, though, if one assumes that the final *n* of *(m)an* has become so integrated that it cannot be construed as the equivalent of the genitive *n* ‘of’ anymore, which could exempt *(m)an* from the necessity to be followed by the annexed state. Furthermore, the following observation by Aikhenvald & Militarev (1991:232) may be relevant here. Aikhenvald & Militarev point out that preposed attributive interrogatives differ from other preposed attributes in that in some Berber languages they allow for the head noun to be in the free state. On the one hand, the variation can be with the same interrogatives between different languages, e.g. Snous *mani argaz* ‘which man?’ (free state) vs. Tamazight *mani uyis* ‘which horse?’ (annexed state) or Izayan Tamazight *matta ayis* ‘which horse?’ (free state) vs. Tarifit *matta uyis* ‘which horse?’ (annexed state).⁴⁴ On the other hand, the variation can be between different interrogatives within one language, e.g. in Tamazight *mani uyis* ‘which horse?’ (annexed state) vs. *matta ayis* ‘which horse?’ (free state). The use of the free state with the preposed attributive interrogatives may be interpreted as an

similar looking demonstrative roots in Semitic languages discussed in Barth (1913:100-103).

⁴³ According to Aikhenvald (1995a:44-45), the “free state” (or “accusative” in her analysis) is typically used in the following functions in Berber: (i) for marking direct objects, (ii) as a citation form, (iii) to code heads in head-dependent constructions (nouns in [N + Adj] and [N + Adp] constructions, the “possessor” [sic! – should be possessed] in genitive constructions), (iv) to code nouns in predicative function or with an existential predicative particle *d*, (v) to code fronted topicalized (and apparently, focalized) constituents. The “annexed state” (or “nominative” in her analysis) is used: (i) to mark non-topicalized (and apparently, non-focalized) subjects, (ii) to mark the dependent in head-dependent constructions (nouns in [Adp + N] and [Num + N] constructions, possessor in possessive constructions).

⁴⁴ Note that *mani* also means ‘where?’. For non-attributive uses of *matta* see Section III.2.3.2.2.5.

indication of the appositional origin of such a construction, i.e. ‘which one, the horse?’ → ‘which horse?’.

2.3.2.2.5 The interrogatives *matta*, *matak*, *ta’k(k)*, and the like

The interrogative *matta* (and the like) is found primarily in Northern Berber languages and often means ‘what?’, as in Tamezret example (70). The interrogative *matta* (and the like) is also very common in attributive use as ‘which [N]?, what (kind of) [N]?’ (see Section III.2.3.2.2.4 for some examples).

Tamezret (Northern Berber, “continuous” Zenati, East Zenati; Tunisia)

- (70) *d mættæ i tægid?*
 it.is IPW DEM 2SG.do.PST.2SG
 ‘What did you do? (lit. ‘What is it that you did?’)’ (Ben Mamou 2005b:8)

Furthermore, this interrogative shows a particular preference for non-verbal clauses, where it is used either in combination with an independent personal pronoun or with a nominal in the so-called “annexed state” (cf. Section III.2.3.2.2.4), as in Ait Seghrouchen (Bentolila 1981:253-254) or Ait Ndir (Penchoen 1973:81). In Ait Seghrouchen, *matta* seems to be applied to persons only in KIND-questions (71), because Bentolila (1981) consistently glosses *matta* and translates questions containing it with French *que*, *quoi* and *qu’est-ce que* ‘what?’.

Ait Seghrouchen (Northern Berber, Tamazight of Atlas or “continuous” Zenati; Morocco)

- (71) a. *awa, matta šk?*
 hey.2SG.M IPW 2SG.M
 ‘Hey, what are you?’ (in the French original as ‘qu’es-tu?’, Bentolila 1981:253)
- b. *šbr, matta uydin?*
 wait IPW AS.that.one
 ‘Wait! What’s that?’ (in the French original as ‘attends! qu’est-ce que c’est que ça?’, Bentolila 1981:254)
- c. *matta uryaz-u?*
 IPW AS.man-this
 ‘What is this man?’ (in the French original as ‘qu’est-ce que c’est que cet homme?’, Bentolila 1981:254)
- d. *matta (n) tɛyyalin-u?*
 IPW GEN AS.this.one
 ‘What are these women?’ (in the French original as ‘qu’est-ce que c’est que ces femmes?’, Bentolila 1981:254)

In Ait Ndir, however, it appears to be possible to use *matta* as both ‘who?’ and ‘what?’ depending on the pronoun or the nominal it is used with, as in (72). Penchoen (1973) consistently glosses *matta* as ‘what?’ but translates it as ‘who?’, when the question is about a person, or ‘what?’, when the question is about a thing.

Ait Ndir (Penchoen 1973:81)

- (72) a. *matta natta?*
 IPW 3SG.M
 ‘Who is he?’
- b. *matta ntta aryaz-ad?*
 IPW he FS.man-this
 ‘Who is this man?’
- c. *matta uryaz-ad?*
 IPW AS.man-this
 ‘Who is this man?’
- d. *matta uya?*
 IPW AS.this.one
 ‘What is this?’

Note that the nature of the syntactic relation between *matta* and the nominal in the annexed state is rather ambiguous and, as Bentolila (1981:253) points out, difficult to analyze in a straightforward way. On the one hand, *matta* can be analyzed as the predicate with the nominal functioning as a postposed subject. The latter is then regularly marked by the annexed state (cf. Section III.2.3.2.2.4), comparably to (73a) vs. (73b). Under this analysis, the construction will be the same as when *matta* is used with a pronominal.

Ait Seghrouchen (Bentolila 1981:255)

- (73) a. *i-rah uryaz*
 3SG.M.SBJ-go.PST AS.man
 ‘The man has gone’
- b. *aryaz i-rah*
 FS.man 3SG.M.SBJ-go.PST
 ‘The man, he has gone’

On the other hand, the possibility of using an overt genitive marker *n* with a feminine nominal in (71d) suggests that, in fact, the construction here is elliptical, viz. [[*what of man*] ...?] > [[*which/what man*] ...?] implying *Who is this man?*. Thus, Penchoen (1973:81) remarks about the construction of Ait Ndir examples (72c-d), that it “seems to have been created secondarily from elliptical

questions referring to the preceding context and in which the noun is a noun complement of *matta*: *matta uryaz?* ‘what of man > which man?’⁴⁵

I have a strong impression that examples like (71a) and (72a), where *matta* is used in a non-verbal clause together with a pronominal, reflect the original usage of *matta*, while its attributive and simple pronominal usages are later extensions (cf. Section III.2.3.2.2.4 and example (70) above). It would be more difficult to explain the reverse development because then we will have to explain why *matta* has become restricted only to this particular context. Furthermore, the proposed direction of the evolution of *matta* readily suggests a plausible etymology of this interrogative, which explains the observed restrictions on its use and can account for the existence of such forms as the Zenaga *matak* ‘which?, what (kind of)?’ (Basset 1909:25) and *ta’k(k)* ‘who?, what?, which one?’ (Catherine Taine-Cheikh, p.c.). I believe that the interrogative *matta* goes back to a combination of an interrogative *ma(d)*, as described in Section III.2.3.2.2.2, and a presentative predicate (or particle) *ta* ‘there [X] is (= French *voilà*)’, such as, for instance, the presentative *ta* found in Figuig (Kossmann 1997:307-309).⁴⁶ That is, the original construction literally meant something like ‘Who/what is that, there he is’. The form *mad* would account for the forms like *matta* with a geminate/tense voiceless dental stop through an assimilation process, very common in Berber languages, whereas the form *ma* would account for the forms with a single *t*, like those of Zenaga. Berber presentative predicates usually combine within one clause only with nominals and pronominals and sometimes only with pronominals, as in the case of the *ta* presentative of Figuig, which is used only with independent plural pronouns (Kossmann 1997:307), as in (74).

Figuig (Northern Berber, Zenati “of oases”; Morocco; Kossmann 1997:307)

- (74) *ta itnin*
 be.there 3PL.M
 ‘There they (M) are’

In various (primarily Northern) Berber languages, presentative predicates can also combine with encliticized direct object personal pronouns instead of the independent personal pronouns, as in (75-76).⁴⁷

⁴⁵ Penchoen’s remark also seems to imply that in Ait Ndir the overt genitive marker *n* is not possible in this construction.

⁴⁶ Maarten Kossmann (p.c.) has suggested that alternatively *matta* may go back to a combination of an interrogative *ma* and the third person singular masculine pronoun *nətta*, which can be compared to the similar pattern common in Arabic (see Section III.2.3.4.1.2). This hypothesis is quite plausible as such. However, unlike the hypothesis involving a presentative predicate, it cannot explain the particular preference of *matta* for non-verbal clauses.

⁴⁷ Aikhenvald (1995a) analyzes this as O=S₀ cross-referencing and uses it as evidence for

Tachelhit (Northern Berber, Atlas Berber; Morocco)

(75) *ha-t-id*

be.here-3SG.M.OBJ-here

‘Here he is’ (Stumme 1899:149 via Aikhenvald 1995a:62)

Ouargla (Northern Berber, Zenati “of oases”; Algeria)

(76) *hak-iy*

be.here-1SG.M.OBJ

‘Here I am’ (Biarnay 1908:149 via Aikhenvald 1995a:62)

Remarkably, presentative predicates often end in *-k(a)*, as illustrated in (77).

(77) *Some Berber presentative predicates ending in -k(V) (based on Aikhenvald 1995a)*⁴⁸

hak ‘be here’ Ouargla (Northern Berber, Zenati “of oases”; Algeria)

‘be there’ Tamazight (Northern Berber, Atlas; Morocco)

aka ‘be this way’ Kabyle (Northern Berber; Algeria)

akka ‘be here’ Rif (Northern Berber, “continuous” Zenati; Morocco)

lak ‘be here’ Mzab (Northern Berber, Zenati “of oases”; Algeria)

nak ‘be here’ Ghadamès (Eastern or Northern Berber; Libya)

This may explain the presence of *-k* in the Zenaga forms *matak* ‘which?, what (kind of)?’ (Basset 1909:25) and *ta’k(k)* ‘who?, what?, which one?’ (Catherine Taine-Cheikh, p.c.). The same *k* also seems to show up in the Ahaggar Tuareg interrogative *manek* ‘which one?’, which, not surprisingly, cannot be used attributively and is regularly marked with the third singular direct object marker *-t* subsequently fusing in *manekk* (Karl-G. Prasse, p.c.), as in (78).⁴⁹

reconstructing a split-ergative alignment of Proto-Berber, with the split being conditioned by the semantics of predicates.

⁴⁸ Maarten Kossmann (p.c.) has suggested that the presentatives *ha* and *hak* may be loans from Arabic. Furthermore, he has pointed out to me that the Rif presentative is *aqqa* rather than *akka* and it is most likely to go back to the verb *qqal* ‘to look’.

⁴⁹ The interrogative *manek* clearly results from the fusion of *mani* ‘where?’ with a presentative **(h)ak(a)*. This may also explain the vowel *e* in *manek* as a result of coalescence of *-i* and *(h)a-*. A similar explanation can be provided for *k* in Malian Tuareg ‘where?, which?’ interrogatives *əndek* (most dialects), *əndəke* (an optional variant in the Kal Ansar variety of the Timbuktu area), and *ənde* (Kidal area dialect; Heath 2005:653). The structure of the first part, *əndV*, appears to be similar to that of the Kabyle interrogative *anda* ‘where?’ (Hamouma 1987:70) discussed in Section III.2.3.2.2.4. The final *-e* may be the *e* allomorph of the 3SG.M direct object enclitic, which would then be similar in origin to the second *-k* in Ahaggar Tuareg

Ahaggar Tuareg

- (78) *manek-k aləs wa tənăyăd ənğom?*
 IPW-3SG.M.OBJ man DEM.M.SG 2SG.SBJ.saw previously
 ‘Which man (among those present now) did you see previously? (lit.:
 ‘Which is he (among those present now), the man you saw previously?’)’
 (Karl-G. Prasse, p.c.)

The latter form *manekk* also provides an indication for the source of the geminate/tense variant of the final stop in Zenaga *ta’k(k)*. Ultimately, the Zenaga interrogative predicate *mahək* ‘where?’ reported by Nicolas (1953), (79), appears to provide an exact structural equivalent to the forms like *matak* (only the presentatives involved are different).

Zenaga (Western Berber; Mauritania; Nicolas 1953:44)

- (79) a. *ma’hək-ki:?*
 IPW-2SG.M.OBJ
 b. *mahək kukʰ?*
 IPW 2SG.M
 ‘Where are you?’

As far as the *-k(a)* of presentatives is concerned, it seems to be cognate to “the deictic and asseverative particle *ka-*, ‘thus’, ‘so’, ‘truly’, [...] common to Semitic languages and to Ancient Egyptian”, which also “plays an important role in the formation of demonstratives” (Lipiński 1997:473).

I would like to conclude this Section by presenting the system of interrogative pronominals of Zenaga of Idab Lahsab, as described by Catherine Taine-Cheikh (p.c.). A summary is provided in (80). The system of Zenaga of Idab Lahsab interrogative pronominals is based on two interrogative elements, *m* and *ta’k(k)*. The first one, *m*, is restricted to use with the human masculine singular proximal demonstrative pronominal *ăd* ‘this’, the resulting meaning of this collocation being *m-ăd* ‘who (is this)?’, as in (81).

Zenaga of Idab Lahsan (Catherine Taine-Cheikh, p.c.)

- (81) *m-ăd y-ūdiž-ăn?*
 IPW-this.M.SG.HUM 3SG.M.SBJ.lie.down.PFV-PTCP.SG
 ‘Who lay down?/ Who is lying? (the original French gloss: *Qui (s’)est couché?*)’

manekk. This *e* allomorph of the 3SG.M direct object enclitic is used “after stem-final *a* or deletable vowel” (Heath 2005:605, 729). I believe this would be a more plausible explanation of why the final *-e* shows up in these interrogatives rather than bringing it back to the truncation of the noun *e-dægg* ‘place’, as Heath (2005:653) cites his informant to suggest.

(80) *Idab Lahsan Zenaga interrogative pronominals (Catherine Taine-Cheikh, p.c.)*

<i>m</i>	<i>äḍ</i>	‘Who is this?’	} (‘...that...’ [Relative Clause])
	this.HUM.M.SG		
<i>ta’k(k)</i>	<i>nəttä</i>	‘Who/what is he/it?’	
	3SG.M (HUM or NON<HUM>)		
<i>ta’k(k)</i>	<i>äḍ</i>	‘Who is this?’	
	this.HUM.M.SG		
	<i>äyḍ</i>	‘What is this?’	
	this.NON<HUM>.M.SG		
<i>ta’k(k)</i>	<i>äḍ</i>	‘Who is that?’	} +
	that.HUM.M.SG		
	<i>iḍ</i>	‘What is that?’	
	that.NON<HUM>.M.SG		
<i>ta’k(k)</i>	<i>ān</i>	‘Who/what is that?’	
	that.M.SG (HUM or NON<HUM>)		
<i>ta’k(k)</i>	N	‘What is the N?’	

The second one, *ta’k(k)*, can be combined with a third person pronoun, such as the singular masculine form *nəttä*. In its reference, *nəttä* is not restricted to humans, but as far as I can judge from Zenaga examples in Faidherbe (1877), it is somewhat more likely to find *ta’k(k) nəttä* as ‘who is he?’. Furthermore, *ta’k(k)* can be combined with one of the demonstrative pronominals, some of which have a human reference, some non-human, and one can have both, which results in ‘who is this/that?’, ‘what is this/that?’ and ‘who/what is that?’ questions respectively, cf. (80). Finally, *ta’k(k)* can be combined with a noun, as in (82). I have examples only with non-human nouns, though.

Zenaga of Idab Lahsan (Catherine Taine-Cheikh, p.c.)

- (82) *ta’kk ta’kšah?*
 IPW problem
 ‘What’s the problem?’

All the combinations of *m* and *ta’k(k)* with nouns and demonstrative and personal pronouns can be used as questions on their own, cf. (80), or they can be followed by a (relative) clause, as in other Berber languages (cf. Section III.2.3.2.2.1), (83-84).

Zenaga of Idab Lahsan (Catherine Taine-Cheikh, p.c.)

- (83) *ta'kk äyḍ* *t-äskär-äḍ?*
 IPW this.NON<HUM>.M.SG 2SG.SBJ-do.PFV-2SG.SBJ
 'What have you done?'
- (84) a. *ta'kk äḍ* *y-ūḍiḥ-än?*
 IPW this.HUM.M.SG 3SG.M.SBJ-lie.down.PFV-PTCP.SG
 'Who is lying?' (the original French gloss: *Quel [est] celui qui [est] couché?*)
- b. *ta'kk t-äḍ* *t-ūḍiḥ-än?*
 IPW F-this.HUM.SG 3SG.F.SBJ-lie.down.PFV-PTCP.SG
 'Who (is this woman who) is lying?' (the original French gloss: *Quelle [est] celle qui [est] couchée?*)

The masculine form of the demonstrative pronominals seems to be the default option, whereas the feminine form, distinguished by adding the prefix *t-*, (84b), seems to be restricted only to 'Who is she?' questions about women and to 'Who/what is the N that?' questions, where N is of the feminine gender.

2.3.2.3 Berber languages: summary

Various Berber languages appear to have at least one 'who?, what?' interrogative pronominal, sometimes more. The fact that languages with interrogative pronominals used indiscriminately as 'who?' or 'what?' can be found in all (but one?) extant major branches of Berber languages and that they are not always the same may be indicative of some structural inclination of Berber languages for developing such interrogatives. This tendency also seems to be a rather stable one, considering that genetically the Berber family is rather deep.⁵⁰

Next to this 'who?, what?' interrogative there is also often a different interrogative for 'who?', while a parallel different interrogative for 'what?' is often lacking. In all probability, this multitude of forms and the frequent occurrence of 'who?, what?' interrogatives are primarily due to the fact that Berber interrogative pronominals are not single morphemes but constructions consisting of a general interrogative root 'who?, what?, which one?' (normally, *m*) and a demonstrative pronominal that typically introduces a relative clause. Structurally, this construction is similar to the one used for the purpose of focalization and can be analyzed as clefting. The kind of demonstrative pronominal involved determines whether the distinction between 'who?' and 'what?' is made.

⁵⁰ It is, for instance, believed to be substantially deeper than the Germanic branch of Indo-European (see Aikhenvald & Militarev 1991:149).

In general, it appears that the *i* based interrogative pronominals (as presented in Section III.2.3.2.2.3) tend to function as exclusively human interrogative pronominals. This is particularly so in the case of *wi*, originally a masculine demonstrative pronominal, but somewhat less universal in the case of *mi*. In the case of the *a* based interrogatives (as presented in Sections III.2.3.2.2.2 and III.2.3.2.2.5), the situation is different, but it is not simply the reverse. On the one hand, there is a rather strong inclination for these interrogatives to function as non-human interrogative pronominals. This is clearly due to the gender-number neutral nature of the demonstrative pronominals involved, which tend to refer to things, actions, etc. (see, for instance, a remark on the Ait Ndir demonstrative pronominal root *ay* in Section III.2.3.2.2.1). But on the other hand the tendency for these interrogative pronominals to develop the ‘who?, what?’ usage is also very much present. This must be primarily due to the particular inclination of the Berber languages to neutralize gender-number distinctions in the case of focalization.

As mentioned in Section III.2.3.2.2.4, sometimes we find similar interrogatives with and without the initial *m*. This raises the question whether the *m* was originally present and disappeared later in some cases or whether it was added later. Given the widespread occurrence of the *m*-initial interrogatives in both Berber and other branches of Afro-Asiatic (cf., for instance, Diakonoff 1965:75 [1988], Ehret 1995:301), the first hypothesis seems to be more plausible. At least in Berber, the initial *m* appears to have long remained a distinct part of a larger construction, which has probably made its occasional ellipsis easier than if it were already fully integrated into the interrogative, like for instance *wh-* of English interrogatives. (Note in this respect that deletion of a morpheme-initial nasal does not seem to be a normal phonological development in Berber, and Afro-Asiatic in general).

The variants without the initial *m* could have also developed not through the ellipsis of *m* but as conventionalization of independently used headless relative clauses, when accompanied by an interrogative intonation. That is, the development may have been from [*Among the people here, the man/ the one*] *who/which is the thief (is)?...* → *Who/Which one is the thief?*⁵¹ If something like this happens, one would normally expect the resulting interrogatives to mean ‘which one?’. This is what is to be expected, because a relative clause normally restricts the possible referent(s) of its antecedent (e.g., *the man who comes here every day*) and, comparably, the interrogative ‘which one?’ asks to choose a referent(s) from a set of possible referents (e.g., [*Among those men,*] *which man/ which one comes here every day?*). Taking into consideration the typologically

⁵¹ A direct development from a demonstrative to an *indefinite* pronominal (and later into an interrogative pronominal), without any intermediate stage like that of a relative pronominal, is semantically somewhat difficult to conceive.

common link between ‘which one?’ and ‘who?’, ‘who?’ would be a further probable development here. In fact, this is exactly what appears to be the case in Berber languages with the interrogatives like *an* ‘which one?’ (Section III.2.3.2.2.4) and *wi* ‘who?’ (Section III.2.3.2.2.3), which are the common *m*-less interrogative pronominals. That the interrogative usage here is likely to represent a secondary development was also demonstrated in Section III.2.3.2.2.3.

2.3.3 *Egyptian*⁵²

2.3.3.1 Introduction

The Egyptian branch of Afro-Asiatic consists of a single language which most likely became extinct in the 16th century. Nevertheless, Egyptian can boast almost five millennia of recorded history, which provide us with a remarkable wealth of linguistic data. As any other natural language, Egyptian was of course constantly changing. For convenience sake, the history of the Egyptian language is divided in five phases and each phase is usually considered as a distinct but related language. Figure 4 is intended to give an idea of how different stages relate to each other. Various sources may differ as to the exact dates and sometimes dynasties, but the overall division in five stages as presented here seems to be universally accepted.

Various interrogative pronominals used both as ‘who?’ and as ‘what?’ can be found in Egyptian all along most of its recorded history, although they usually tend to become restricted in one of the two meanings. By the Coptic period, the system of Egyptian interrogative pronominals becomes substantially restructured, in all probability under Greek influence. I will discuss Old and Middle Egyptian interrogatives together in Section III.2.3.3.2 because the two idioms are generally believed to be quite similar. Furthermore, Old Egyptian data on interrogatives are considerably less informative because most sources of this period represent the formal register. Late Egyptian will be considered in Section III.2.3.3.3, Demotic in Section III.2.3.3.4 and Coptic in Section III.2.3.3.5.

Before proceeding further, one caveat should be noted with regard to the pre-Coptic Egyptian writing systems and the strength of the claims that they allow us to make about Egyptian. Although hieroglyphic script is predominantly phonetic and only partially ideographic, its usability for linguistic research is seriously hampered by the fact that it is almost exclusively consonantal.⁵³

⁵² I am grateful to Jean Winand for his comments on the present section.




⁵³ Presumably, the signs  (G43) and  (M17), usually transliterated as *w* and *i~j* respectively, are the only exceptions. These signs, especially , are believed to represent not only approximants, which phonologic value their transliterations indicate, but also “a vowel that was present in presumably an unexpected position, without necessarily indicating what vowel was

Figure 4. Linguistic stages of Egyptian (based on Gardiner 2001:5 [1957] & Callender 1975:1)

Old Egyptian	The language of Dynasties I-VIII, ca. 3200 B.C. to 2200 B.C. The sources of this stage are mainly official or other formal documents, such as funerary formulae and tomb inscriptions, including some biographical texts.
Middle Egyptian	Also known as “ Classical Egyptian ”, the language of Dynasties IX-XVII, ca. 2200 B.C. to 1600 B.C. In structure and vocabulary it is very close to Old Egyptian. In its later form, it survived down to the Greek-Roman period for formal purposes.
Late Egyptian	The language of Dynasties XVIII-XXIV, ca. 1600 B.C. to 700 B.C., used for informal and semi-formal purposes. It is radically different in structure from Middle Egyptian, but there are few texts without any admixture of the “classical” language. Reversely, the texts in “classical” language produced after the Middle Egyptian period are inevitably contaminated with elements from later stages.
Demotic	This term is loosely applied to the language written in a cursive script, known as “demotic”, which is ultimately derived from the hieroglyphic writing system. It is similar in structure to Late Egyptian and was in use from Dynasty XXV to late Roman times, ca. 700 B.C. to 300 A.D.
Coptic	The form of Egyptian, as written in Coptic script and still used for liturgical purposes in the Coptic Church. It ceased to be a living language somewhere in the 16 th century. In structure it is basically akin to Demotic and Late Egyptian but heavily-influenced by Greek.

Generally speaking, vowels (and their quality) become marked with a certain regularity only in the later demotic variant of the hieroglyphic script (Lexa 1947-1951:9). However, it is only with the Coptic script based on the Greek alphabet that vowels become fully specified in writing. Therefore, when speaking about pre-Coptic Egyptian, a chance will always remain that the interrogative

present” (Callender 1975:3). A vocalic interpretation appears to be most plausible word-initially. The double sign 𐀀 (or 𐀁), usually transliterated as *y*, is hardly found as initial letter and may represent just a graphic variation of its singleton counterpart (Callender 1975:3, 5).

If not mentioned otherwise, Egyptian examples here are given in the traditional Egyptological transliteration, as can be found in Gardiner (2001:27 [1957]).

pronominals presumed to be used both as ‘who?’ and as ‘what?’ are in fact two interrogatives, ‘who?’ and ‘what?’, differing only in their vocalization.

Nevertheless, the probability of the latter situation seems to be rather low. First, although homographs were usually further distinguished by means of additional signs, the so-called “determinatives”, this never seems to be done for interrogative pronominals. A given interrogative pronominal is always rendered with the same sign(s) whether it means ‘who?’ or ‘what?’, which strongly suggests that it is always one and the same interrogative. At the very least, this seems to suggest that even if one writing does stand for more than one interrogative differing only in their vowels, any difference possibly expressed this way was not lexical, ‘who?’ vs. ‘what?’, but grammatical of some kind, e.g., difference in case, the so-called “state”, focus, etc. Second, it appears that the agreement patterns of a presumed ‘who?, what?’ interrogative pronominal do not depend on whether it means ‘who?’ or ‘what?’ (see Section III.2.3.3.2.1). Third, more than one unrelated ‘who?, what?’ interrogatives are recorded, which is indicative of a certain recurrent pattern in the development of Egyptian interrogative pronominals. Furthermore, for at least one of these interrogatives, *pw* (more common in combination with a particle as *pw-tr/ pw-ti/ ptr/ pty/ pt*), Gardiner (2001:406 [1957]) has suggested a single source, a demonstrative *pw* ‘DEM.M.SG’, which implies that the writings representing the respective interrogative are also likely to have only one vocalization for both ‘who?’ and ‘what?’. Finally, additional indirect evidence in favour of the existence of ‘who?, what?’ interrogatives in Egyptian is provided by the fact that such interrogatives are not uncommon in the two other geographically, and probably genetically, most close branches of Afro-Asiatic, which are Berber and Semitic (see Sections III.2.3.2 and III.2.3.4).

2.3.3.2 Old and Middle Egyptian

The Old and Middle Egyptian interrogative pronominals are summarized in (85). Transliteration variants will be discussed in the sections on the respective interrogatives below. In what follows, I base myself on Edel (1955, 1964) for Old Egyptian and Allen (2000), Callender (1975), Gardiner (2001 [1957]), Korostovtsev (1991) and Loprieno (1996) for Middle Egyptian. The interrogative *iḥ* is not reported for Old Egyptian and it becomes frequent only by the Late Egyptian period. Therefore, although it is introduced here, it will be considered in more detail in Section III.2.3.3.2.

(85) *Old and Middle Egyptian interrogative pronominals (based on Allen 2000, Edel 1955, 1964, Callender 1975, Gardiner 2001 [1957], Loprieno 1996)*

a. <i>m, mi</i>		‘who?, what?’
b. <i>pw</i>		‘who?, what?’
<i>pw-tr</i>		
<i>pw-ti</i>		
<i>p-tr</i>		
<i>p-ty (p-ti)</i>		
<i>p-t</i>		
c. <i>iš-st</i>		‘what?’ (unlikely also ‘who?’)
<i>iš-sy</i>		
d. <i>ih</i>		‘what?’
e. <i>zy (zi), sy (si)</i>		‘which [N]?, what (kind of)
<i>z, s</i>		[N]?, ‘which one?, who?, what?’

With the exception of *zy*, the interrogatives summarized in (85) are not used attributively. When they are translated as adjectival ‘what (kind of)?, which?’, they are used either in a genitive construction with *n(y)* ‘of’ or with a preposition *m* ‘as’ (Gardiner 2001:406 [1957]), as in (86).

Middle Egyptian

(86) *n m n ntr?*
to IPW of god

‘to what god? (lit.: ‘to whom of/as god?’) (Gardiner 2001:406 [1957])⁵⁴

According to Gardiner (2001:406 [1957]), such usage is rare, though. As to the attributive usage of *zy*, as in (87), Allen (2000:55) considers *zy* to be “the first noun of a direct genitive”, that is a genitive construction with a genitive marker *n(y)* ‘of’ (sometimes also called *nota genitivi*), whereas Gardiner (2001:407 [1957]) suggests that the noun following *zy* is “in apposition to it”.

⁵⁴ As a rule, glosses in Egyptian and Coptic examples have been added by me, except in examples from Callender (1975) where the original glosses have just been sometimes modified for the sake of uniformity.

Middle Egyptian

- (87) *zy/s(y) w3t?*
 IPW/IPW path
 ‘which path?’ (Allen 2000:55), ‘[on] what road?’ (Gardiner 2001:407 [1957])

Similarly to constituent questions in Berber languages (Section III.2.3.2.2.1), Egyptian constituent questions are regularly built up as focalization constructions, which can be construed as a kind of clefts, although as Callender (1975:91) remarks, “the term ‘clefting’ is as yet not widely used in Egyptological studies”. The clefting implies that interrogative pronominals function as predicates and that as a rule, the verb is converted into an “(active or passive) participle” or a “relative form” (for more details on these terms, see for instance Gardiner 2001 [1957]). Consider, for instance, the following characterizations of the Egyptian constituent questions:

Egyptian requires that all non-rhetorical questions containing interrogative words such as ‘who’, ‘what’ or ‘where’ be clefted. The type of clefting depends on the status of the interrogative word, with the one constraint that interrogative words that form parts of adverbial phrases cannot be clefted as such. [...] Adverbial interrogatives must be clefted with the manner nominalization.

Callender (1975:96, 98)

As a general rule, interrogative pronouns behave like focalized subjects or objects of nominal predicates. The focalized subject pronoun *(i)n-m* “who?” [...] occupies the position of the independent pronoun in a specifying pattern [as in the first sentence in (88), which can be compared to (89)].

Loprieno (1996:121)

Middle Egyptian

- (88) *(i)n m tr tw? ntk zi?*
 SBJ.FOC(= it.is) IPW actually 2SG.M.DEP 2SG.M.INDEP IPW
 ‘Who are you then? Who are you?’ (Loprieno 1996:121), ‘Who pray art thou?’ (Gardiner 2001:405-406 [1957])
- (89) *ntf hrw*
 3SG.M.INDEP Horus
 ‘He is Horus’ (Loprieno 1996:104) or rather ‘HE is Horus’ (with emphasis on the pronoun, Gardiner 2001:103 [1957])

It should be mentioned that fronting, as in (90), is a common but not indispensable attribute of focalization techniques in Egyptian.

Middle Egyptian

- (90) *p-tr* *dd-t* *n-i* *nb-i*
 IPW-actually say.IPFV.REL-F to-1SG.SUF lord-1SG.SUF
 ‘What does my lord say to me? (lit.: ‘What is actually that which my lord says to me?’)’ (Gardiner 2001:407 [1957])⁵⁵

In fact, interrogative pro-words also often remain *in situ* (Gardiner 2001:404-405 [1957]), especially in the case of interrogative pronominals that are part of a prepositional phrase, as in (91).

Middle Egyptian

- (91) *hpr-n* *mdw m s3wy* *r(3) mdw n m?*
 become.PFV.(M)-of ten as two-thirds one tenth of IPW
 ‘Of what is ten the 23/30 part? (lit.: ‘Ten has become as two-thirds (and) one-tenth of what?’)’ (Gardiner 2001:405 [1957], Callender 1975:98)⁵⁶

Different positions of interrogative pro-words can be illustrated by (92a) and (92b) from Middle Egyptian and (93a) and (93b) from Old Egyptian.⁵⁷

Middle Egyptian

- (92) a. *tw* *m tr?*
 2SG.M.INDEP IPW actually
 ‘Who are you then?’ (Loprieno 1996:121), ‘Who pray art thou?’
 Gardiner 2001:405 [1957])⁵⁸
 b. *m tr tw?*
 IPW actually 2SG.M.DEP
 ‘Who pray art thou?’ (Gardiner 2001:405 [1957])⁵⁹

⁵⁵ The feminine agreement on the verb in this example refers to the inanimate ‘what?’. See below in Sections III.2.3.3.2.1 and III.2.3.3.2.3 for more details.

⁵⁶ The verb form glossed here as “PFV-of” is usually referred to in Egyptology as “*šdm.n.f* (or *sdm.n.f*) form” (Gardiner 2001:55-56, 326-328 [1957]), while Callender (1975) calls it “past manner nominalization”. It can be used in a rather broad range of functions, mostly with past reference (for more details, see Gardiner 2001:329-334 [1957]). For convenience sake, I choose the gloss PFV, which represents one of the possibilities to gloss it with a more practical label (Jean Winand, p.c.). Another example of the “*šdm.n.f* form” can be found in (93).

⁵⁷ According to Jean Winand (p.c.), different placement of the interrogative is associated with some semantic differences.

⁵⁸ The form *tw* is an archaic form of the 2.SG.M.INDEP pronoun *ntk*, as found in (88).

⁵⁹ Here the dependent form of the pronoun *tw* is used instead of the independent form *tw* because independent pronouns “almost always stand at the beginning of the sentence” (Gardiner 2001:53 [1957]).

Old Egyptian

- (93) a. *iri-n-i m?*
do.PFV.M-of-1SG.SUF IPW
‘What have I done? (lit.: ‘The one that has been done by me (is) what?’)’ (Edel 1964:517)
- b. *m hwi-n-k?*
IPW beat.REL.M-of-2SG.M.SUF
‘Whom have you beaten? (lit.: ‘(It is) who which has been beaten by you?’)’ (Edel 1964:516)

2.3.3.2.1 The interrogative *m(i)*

The interrogative *m* means both ‘who?’, as in (86, 88, 92, 93b), or ‘what?’, as in (91, 93a, 94). Morphosyntactically, *m* behaves rather like a “dependent” pronoun in that it “is used mostly after other words” (Allen 2000:54), such as, for instance, a preposition or a genitive marker, as in (86), (91) and (94), a subject focus marker *in*, as in (88), a verbal form, as in (93a), etc. It very rarely begins a sentence on its own, as in (92b) and (93b).

Middle Egyptian

- (94) a. *m-m?*
with-IPW
‘with what?’ (Gardiner 2001:406 [1957])
- b. *mi-m?*
like-IPW
‘how?’ (Gardiner 2001:406 [1957])
- c. *r-m?*
to-IPW
‘to what purpose?, what for?’ (Gardiner 2001:406 [1957])
- d. *hr-m?*
on-IPW
‘why?, because of what?’ (Gardiner 2001:406 [1957])

When asking about the subject the interrogative *m* is regularly preceded by a subject focus maker *in*. Usually, this combination means ‘who?’, as in (88) and (95).

Middle Egyptian

(95) *in m irr?*

SBJ.FOC IPW do.IPFV.(M)

‘Who makes (acts)?’ (Gardiner 2001:406 [1957])⁶⁰

Already in Middle Egyptian and maybe even in Old Egyptian the combination *in m* tends to fuse in *n-m*, which eventually will result in the Late Egyptian and Coptic interrogative *nim* ‘who?’ (Gardiner 2001:406 [1957], Edel 1964:515, Till 1986:102 [1966]). Note, however, that *in m* is also attested in the meaning ‘what?’, as in (96).

Middle Egyptian

(96) *in m dd sw?*

SBJ.FOC IPW say.IPFV.(M) 3SG.M.DEP

‘[A:] What expresses it? [B: Twenty expresses it].’ (Gardiner 2001:405 [1957], Callender 1975:97)

Interestingly, according to Callender (1975:92), in declarative sentences, *in* tends to mark subject focus only with specific (or definite?) subjects, whereas a cleft construction based on the “copula” *pw* (originally, a demonstrative ‘DEM.M.SG’) is used when the subject is non-specific (or indefinite?). In Loprieno’s (1996:104) analysis, the difference is here between cleft and pseudocleft, respectively, as illustrated in (97a) and (97b).

Middle Egyptian

(97) a. *in rmt hzi wi*

SBJ.FOC man praise.PFV.(M) 1SG.DEP

‘It is the man who praised me’ (Loprieno 1996:104)

b. *rmt pw hzy-n-f?*

man COP/DEM praise.PFV.(M)-of-3SG.M.SUF


‘The one whom he praised is a man’ (Loprieno 1996:104), or as Callender (1975) would translate it, ‘It was a man whom he praised’ or ‘He praised A MAN’

Note that *pw* in (97b) can also be analyzed as a relative pronoun ‘that, who, whom, which’, [*rmt* ‘(It is) a man’]_{Main Clause} [*pw* ‘that’ *hzy-n-f* ‘he praised’]_{Subordinate Clause}, which would be strongly reminiscent of the use of demonstratives in Berber clefts (cf. Section III.2.3.2).

Another point to be mentioned about the interrogative *m* is its agreement pattern. Remarkably, it appears that its agreement pattern does not depend on whether it means ‘who?’ or ‘what?’ or whether it is preceded by *in* or not.












⁶⁰ According to Jean Winand (p.c.), this is a constructed example.











Whenever *m* can control agreement (i.e., especially in combination *in m*), the agreement is masculine, as in (93, 95, 96) above. In fact, this fits rather well with the fact that as mentioned above, this interrogative has already specialized as ‘who?’ by the Late Egyptian period, because in Old and Middle Egyptian “the meaning of the *neuter* [was still] expressed [...] by the feminine”⁶¹ and only by the Late Egyptian period the masculine becomes fully established as the default choice for this purpose (Gardiner 2001:417, 48 [1957]). In this respect, it is instructive to compare agreement patterns used with other interrogatives, such as *išst* (Section III.2.3.3.2.3) or *ptr* (and the like; Section III.2.3.3.2.2), both ‘what?, who?’ but much more common as ‘what?’. Thus, the latter two interrogatives tend to trigger feminine agreement, as in (90) and (107), and generally they both mean ‘what?’ rather than ‘who?’.

Various writings of the interrogative *m* are attested, although they are never used to distinguish between ‘who?’ and ‘what?’ meanings. These writings together with evidence from Coptic and related Afro-Asiatic languages may provide us with a clue to the original phonological form of this interrogative, which was either /m/ or /mi ~ ma/, or perhaps both until a certain moment. In Old Egyptian, the most typical writing seems to be simply the phonogram  *m* (G17)⁶² (Edel 1955:90, 1964:515-517), which would normally imply that this form was realized as /m(V)/. What is important here is the absence of a vowel before *m* and of a second consonant at the end.⁶³ However, already in Old

⁶¹ Perhaps, the feminine was originally preferred in this function because the noun (*i*)*ht* ‘thing(s), something, property’ is feminine (cf. Gardiner 2001:583 [1957]).

⁶² The letter and number represent the codes used in Gardiner’s (2001 [1957]) classification of hieroglyphs, which is traditionally used in Egyptology.

⁶³ Edel (1964:516-517) also reports for Old Egyptian such writings of the interrogative as  (M17-G17) and  (M17-G17-D35), which he characterizes as “remarkable writings” (“auffälligen Schreibungen”), though. Note that in Edel’s examples these writings show up only when the interrogative is in sentence-initial position. In all probability, they represent contractions of *in m* due to assimilation: *im(m)* in the first case and *in(n)* in the second case. The sign  (M17) *i* comes then from *in* , composed of two uniliteral phonograms G17 and N35. The sign  (D35) as an ideogram or a determinative expresses the idea of negation, whereas its phonetic value is transliterated as *n*. An assimilation from *in m* to *im-m* or to *in-n* would have resulted in two identical consonants next to each other and in such cases there was “a strong tendency to write them but once” (Gardiner 2001:52 [1957]). In  (M17-G17), *im(-m)*, only *m*  (G17) would naturally be written. In  (M17-G17-D35), *in(-n)*, however, it was probably considered more important to preserve the sign  (G17) indicating the interrogative in order to avoid possible ambiguity. In this case, the sign  (D35) can be interpreted as a phonetic determiner specifying the reading of the previous sign. It would then be similar to other determinatives used with  (G17) in the interrogative at issue (see further down in the present Section).

Egyptian times and particularly in the Middle Egyptian period the writing of the interrogative is extended with a so-called determinative, usually  (D38) or  (D36), sometimes  (D37). The latter three signs can hardly be semantic determinatives here because then they normally refer to the idea of giving. Apparently, in the case of the interrogative, they function as phonetic determinatives specifying the reading of another sign (group of signs). The sign  (D36) is of little help because it is often used instead of other “forearm”-signs (D37-D44). The signs  (D38) and  (D37) when used as a phonogram have the value *mī* or *m*. Probably, this is why Allen (2000:54) transliterates the interrogative as *mī*. Furthermore, according to Gardiner (2001:469 [1957]), the combinations of signs  (G19) and  (G20), also sometimes used to render this interrogative, originally had the value *mī*. The transliteration *mī* may suggest that by the Middle Egyptian period the interrogative /m(V)/ has developed a new form /mVj(V)/, the two forms being perhaps maintained next to each other. It seems unlikely that the new writing *mī* would have evolved just as an equivalent to the older one *m* as a new means of representing the same phonological form /m(V)/. The reason is that the value of the signs  (D38) and  (D37) as phonograms originates in their usage in the writings of an irregular imperative *imi* ‘give!’ (Gardiner 2001:454 [1957]) and this imperative has survived in Demotic as *my* (and not *mī*) and in Coptic as *mai* /maj/, in some dialects reduced to *ma*, (Vycichl 1984:103), which suggests that the writing *mī* should be read as /mVj(V)/.

If /j/ here results from hiatus resolution after adding a vocalic morpheme /-V/, then the preceding vowel should have been /i/. A hypothetical development /mi/ + /-a/ > /mija/ would not be unnatural. Otherwise, /j(V)/ may have been added in its entirety, originating then probably in a deictic root.⁶⁴ The form /mVj(V)/ can be compared to Ugaritic *my* ‘who?’ (Afro-Asiatic, West Semitic; Syria; data for the period ca. 1400-1190 BC; Pardee 1997:134) and Old Canaanite of El-Amarna letters *miya* ‘who?’ (Afro-Asiatic, West Semitic; data for the period ca. 1350 BC; Segert 1997:178), which would suggest the reading /mij(V)/ for the Ancient Egyptian form. Otherwise, the form /mVj(V)/ can be compared to the Ait Ndir *m-ay* ‘who?, what?’ and similar Northern Berber forms (Afro-Asiatic, Northern Berber; Morocco; cf. Section III.2.3.2.2.2), which would suggest the reading /maj(V)/. In turn, the form *m(V)* can be compared either to Malian Tuareg *mi* ‘who?’ and ‘where? (what?)’ (Afro-Asiatic, Southern Berber; Mali; see Section III.2.3.2.2.3), Ait Ndir *mi* ‘who?, what?’ (Afro-Asiatic, Northern Berber; Morocco; see Section III.2.3.2.2.3), Tigre *mi* ‘what?’ (Afro-Asiatic, West

⁶⁴ For instance, a demonstrative root *i* is indicated by Loprieno (1996:68). For possible Semitic parallels of this demonstrative root see, for instance, Barth (1913:89-91, 115-116, 129-130). Combining interrogative pronominals with “reinforcing” deictics is quite common in other branches of Afro-Asiatic (cf. Sections III.2.3.2 and III.2.3.4.1.2).

Semitic; Eritrea; Raz 1997:448) or Proto-Chadic *mi* ‘what?’ (Frajzyngier 1985:64), suggesting the reading /m(i)/. Otherwise, it can be compared to the Malian Tuareg *ma* ‘what?’ (Heath 2005:651) and Classical Arabic *mā* (< **mah(a)*) ‘what?’ (Afro-Asiatic, Western Semitic; Lipiński 1997:328-329), suggesting the reading /m(a)/. It should be kept in mind, however, that since in most Afro-Asiatic languages the consonantal skeleton of a wordform is usually more important than its vocalic filling for purposes of lexical differentiation,⁶⁵ the vowel before /j/ might have originated just as a kind of filler. In this case, it would have been a front vowel, but its quality was fully determined by the following /j/.⁶⁶

The Coptic reflex of the interrogative *m(i)*, *nim* ‘who?’ (cf. Section III.2.3.3.5), results from a fusion of the subject focus particle *in* with the interrogative *m(i)* (Gardiner 2001:406 [1957], Edel 1964:515, Till 1986:102 [1966]). The focus particle was most likely vocalized as /(?)an(V)/, as discussed in footnote 42 in Section III.2.3.2.2.4. The *i* in the Coptic form *nim* is of little help in establishing the vocalization of the interrogative *m(i)*, since it is either a reflex of the final vowel of the focus particle, which should then be vocalized as /(?)ani/, or an originally epenthetic vowel between the final *-n* of the focus marker and the initial *m-* of the interrogative.⁶⁷

Some vocalized Demotic writings of the reflexes of *in m(i)* may be more helpful. Thus, Spiegelberg (1925:19) reports a writing transliterated as *nime* (next to *nim* and, without vocalization, *nm*). The Demotic final *e* normally results in Coptic *ε* (Lexa 1947-1951:44), also transliterated as *e* and usually assumed to be realized as /ε/ or /ə/. This may point at an earlier reading of *mi* as /ma(j)/ rather than /mi(j)/.

Finally, the Ait Ndir (Afro-Asiatic, Northern Berber; Morocco) interrogative *mi* ‘who?, what?’ discussed in Section III.2.3.2.2.3 shows an intriguing similarity in its morphosyntactic behaviour to the Old and Middle Egyptian interrogative.

⁶⁵ Just recall in this respect the often-cited fact that most scripts developed by speakers of the Afro-Asiatic languages are (predominantly) consonantal.

⁶⁶ This may be compared to the first vowel of the Proto-Semitic third person singular independent pronouns, *u* in the masculine **šū-wa* and *i* the feminine **šī-ya*, which according to Lipiński 1997:303-304, just “corresponded qualitatively to the semivowels *w* and *y* [i.e., /j/]”.

⁶⁷ The fact that there are no traces of /i/ in the Coptic “independent” (or “stressed”) pronouns, which all begin with a reflex of the Egyptian focus particle *in* (cf. Loprieno 1996:64-65), might be indirect evidence in favour of the epenthesis hypothesis. If /i/ is epenthetic, it must have appeared there already during the Middle Egyptian period as some writings given by Gardiner (2001:406 [1957]) may suggest. The choice of /i/ as an epenthetic vowel is not surprising, because for instance in Classical and Modern Standard Arabic /i/ is the “default juncture [i.e., epenthetic] vowel” used to prevent “non-canonical clusters of consonants arising between words in juncture” (Badawi *et al.* 2004:10).

Thus, it is never used on its own and must always be preceded by a preposition (or a genitive marker, which sometimes is also analyzed as a preposition).⁶⁸ Recall in this respect that interrogative *m(i)* is quite similar in that it tends to behave as a “dependent” pronoun in Old and Middle Egyptian. This may also suggest that in early stages of Ancient Egyptian the interrogative was just an enclitic /m/ without any intrinsic vowel.

All the above direct and indirect evidence taken together suggests that the interrogative *m* was read either /m/ or /mi ~ ma/. In the latter case, it is most likely that the two forms /mi/ and /ma/ did not differ lexically, as ‘who?’ vs. ‘what?’, at least not directly so. Rather, as suggested by the “dependent” use of the Old and Middle Egyptian interrogative and its Ait Ndir counterpart, the difference was of a grammatical kind. It might have been a difference in case or the so-called “state” (cf. Section III.2.3.1).⁶⁹ In the course of time, first a demonstrative /j(V)/ has been added and then one form has become generalized.⁷⁰ The Demotic evidence presented above seems to indicate at /ma(j)/ rather than /mi(j)/ as the best candidate for such a generalized form.

2.3.3.2.2 The interrogatives *pw*, *ptr*, *pty*, and the like

Among the various forms of the *p*-interrogatives presented in (85b), the original one is *pw*, all other forms resulting from a combination with a particle *tr/ ty/ ti/ t* ‘actually, forsooth, I wonder’. According to Gardiner (2001:406 [1957]), *pw* stems from the demonstrative *p-w* ‘M.SG-DEM’, which can be compared to the development of the interrogative *wi* ‘who?’ from a masculine demonstrative in various Berber languages (see Section III.2.3.2.2.3). Although the interrogative *pw* is not attested in Old Egyptian texts (Edel 1955:90) and is very rare in Middle Egyptian (Gardiner 2001:407 [1957]), it is believed to be old (“sicherlich als alt anzusehen”, Edel 1955:90). Various combinations of the interrogative *pw* and the particle *tr/ ty/ ti/ t* ‘actually, forsooth, I wonder’ are, on the contrary, quite common, especially the contracted forms, such as *ptr* and *pty*.

As to the meaning of the interrogative *pw*, Edel (1964:518) gives only ‘who?’, as in (98).

⁶⁸ Somewhat similar behaviour is also attested for a rare Geez (Afro-Asiatic, West Semitic, Ethiopian; Ethiopia; data for the period ca. 350-1000 AD) interrogative *mi(:)* ‘what?’, which according to Barth (1913:141-142), is always used as a “proclitic bound to the following word” (“immer proklitisch mit dem nächsten Wort verbunden”).

⁶⁹ See, for instance, Loprieno (1996:65) who speaks of “traces of ergativity, together with other remnants of a full-fledged case system” in Egyptian.

⁷⁰ The generalization of one of the forms did not have to coincide with the loss of such grammatical opposition in the language in general.

Old Egyptian

(98) *pw sw (i)ḥk(i)?*

IPW 3SG.M.DEP enter.IPFV.(M)

‘Who is he who enters?’ (Edel 1964:518, Gardiner 2001:407 [1957])⁷¹

Gardiner (2001:407 [1957]) glosses *pw* as both ‘who?’ and ‘what?’, although without providing any clear example of its use as an interrogative ‘what?’. Therefore, it is not excluded that originally *pw* was used only, or preferably, as ‘who?’. In all probability, this is due to the fact that it originates in a *masculine* demonstrative *pw*. As discussed in Section III.2.3.3.2.1, in Old and Middle Egyptian “the meaning of the *neuter* is expressed [...] by the feminine”, the masculine being avoided in such cases. Only by the Late Egyptian period does the masculine replace the feminine as the default choice for this purpose (Gardiner 2001:417, 48 [1957]). To a certain extent, this shift to the masculine for expressing the meaning of “neuter” can account for the fact that a later, more frequent *ptr* (and the like) form of this interrogative appears to be preferably used as ‘what?’ rather than ‘who?’.⁷²

In addition, the following evolution that the demonstrative *pw* has undergone during the Old Egyptian period (“long before the Middle Kingdom”, Gardiner 2001:104 [1957]) must have played an equally important role here. Thus, *pw* DEM.M.SG came to be used first as “logical subject after logical predicates consisting of a noun [...] as an equivalent for ‘he’, ‘she’, ‘it’ or ‘they’ *invariable in number and gender*” (italics added), as in (99) and (100), and subsequently as a copula linking two nominals, as in (101) and (102), (Gardiner 2001:103-104 [1957]; consider also a possibility of an alternative analysis of *pw* in (97b) as a relative pronoun discussed in Section III.2.3.3.2.1).

Middle Egyptian

(99) *Rḥ p-w*

PROP M.SG-DEM

‘[A: Who is he?, B:] It/He/This is Ra (name of a god)’ (Gardiner 2001:103 [1957])

⁷¹ The initial and the final *ḥ*’s in the verb are put in brackets because they do not appear to be represented in the hieroglyphic writing that Edel gives for this example, nor are they transliterated by Gardiner in an almost identical example provided in his grammar of Middle Egyptian. That two *ḥ*’s should not be present has been confirmed to me by Jean Winand (p.c.), who further comments that in some variants of the text where this example comes from the interrogative is written without the final *w*, as if it were cliticized.

⁷² Usually, this tendency is not explicitly mentioned in the sources but can be deduced from the fact that in the overwhelming majority of examples provided in the sources, *ptr* (and the like) is used as ‘what?’.

- (100) *ḥmt w^b p-w n R^s*
 wife(F) priest M.SG-DEM of PROP
 ‘[A: Who is this Reddjedet?, B:] She is the wife of a priest of Ra (name of a god)’ (Gardiner 2001:103 [1957])
- (101) *dmī pw imnt*
 abode COP west
 ‘The West is an abode’ (Gardiner 2001:104 [1957])
- (102) *mnw pw n s nfrw-f*
 monument COP of man goodness-3SG.M.DEP
 ‘A man’s goodness is his monument’ (Gardiner 2001:104 [1957])

Remarkably, it appears that unlike other Old and Middle Egyptian interrogative pronominals, the interrogatives *pw*, *ptr*, etc. always function as predicates (usually of clefts, but not only, cf. examples (104) and (105) below) and never as objects, nor can they be headed by a preposition. This can be considered as evidence in favour of the link between the interrogative and the demonstrative-cum-copula which was mentioned in the beginning of the present section.

The agreement pattern of the interrogative *pw* is not clear because in the examples I found, such as (98), it is used together with a pronoun so that the agreement pattern on the verb is more likely to be determined by the pronoun rather than by the interrogative. However, for *ptr* (and the like) there are enough examples showing a tendency for the use of the feminine agreement pattern, as in (90) above and (103) here.

Middle Egyptian

- (103) *p-ty h33-t r-f m sšr?*
 IPW-actually descend.IPFV-F to-3SG.M.SUF as corn
 ‘What amount of corn can go into it? (lit.: ‘What is that which descends into it in corn?’)’ (Gardiner 2001:406 [1957], Callender 1975:97)

The use of the feminine here can be compared to the use of the same feminine agreement pattern with the interrogative *išst*, which basically means ‘what?’ (see Section III.2.3.3.2.3).

Examples (104) and (105) further illustrate the use of the interrogative *ptr* (and the like).

Middle Egyptian

- (104) *p-tr rf sw?*
 IPW-actually then 3SG.M.DEP
 ‘Who is he?’ (Gardiner 2001:406 [1957]) or ‘Who is he?, What is it?’ (Callender 1975:97)

- (105) *p-ty* *n-3?*
 IPW-actually N-DEM
 ‘What is this/that?’ (Gardiner 2001:86 [1957], Allen 2000:52)

Note that in these two examples, the interpretation of the interrogative as ‘who?’ or ‘what?’ can be deduced only through the demonstrative or pronoun it is used with. This is reminiscent of the way Zenaga of Idab Lahsan (Afro-Asiatic, Western Berber; Mauritania) interrogative pronominals function, as examined in Section III.2.3.2.2.5 (in particular, cf. (80)).

2.3.3.2.3 The interrogative *išst/išsy*

The interrogative *išst/išsy* means ‘what?’. The possibility that it can also be used as ‘who?’ will be examined later in this section. As far as the difference between the two forms *išst* and *išsy* is concerned, *išst* is the regular form. The transliteration *išsy* is given at one place by Gardiner (2001:407 [1957]) (cf. (107) below). However, for some reason, he specifies that it should still be read as *išst* (Callender 1975:97 gives the same example simply with *išsy*). An etymology of this interrogative proposed by Allen (2000:55) could justify both transliterations and at the same time, explain the preference for the form *išst*.

First, Allen (2000:55) suggests that the *iš-* part of this interrogative is a variant form of another, less frequent interrogative *iḥ* ‘what?’ (cf. Section III.2.3.3.3.2), the latter, according to Gardiner (2001:408 [1957]), going back to a feminine noun (*iḥ-t* ‘thing(s), something, property’ (-*t* is the feminine suffix).⁷³ Second, Allen (2000:55) believes that the *-st* part of the interrogative at issue originates in a third person dependent pronoun *st*. Allen (2000:49) analyzes the dependent pronoun *st* as “neutral in gender and number” and referring to things (“nouns or plurals” other than those designating “living beings (people or gods)”). Other third person dependent pronouns, *sw* ‘he, him’, *š* ‘she, her’ (also transliterated as *sy* in other sources) and *sn* ‘they, them’, are said to refer “for the most part [...] to living beings (people or gods)”. I believe that (i) the combination of the original feminine gender of *iḥ/iš-*, with the already mentioned

⁷³ The link between the word for ‘thing’ and *iš* can be supported by the fact that in Old Egyptian, *iḥt* was written as *išt* before suffixes (Hannig 2003:200-201). However, I am reluctant to accept Allen’s hypothesis that *iš-* of the interrogative *išst/išsy* originates in the interrogative *iḥ* ‘what?’. Rather, *iḥ* ‘what?’ represents a later development from the same source, the noun ‘thing, something’. Note, for instance, that while *išst/išsy* is well attested in Old Egyptian, *iḥ* ‘what?’ is not (cf. Edel 1955:90, 1964:515-518). Furthermore, *iḥ* is also rare in Middle Egyptian becoming common only in Late Egyptian (Gardiner 2001:408 [1957]). Finally, the variation *ḥ* ~ *š* in the word for ‘thing’ is reported for Old Egyptian rather than for later stages.

Old and Middle Egyptian preference for the feminine to be used for expressing “the meaning of the *neuter*” (Gardiner 2001:417, 48 [1957]; see Sections III.2.3.3.2.1-III.2.3.3.2.2) and (ii) the functional distribution between the pronouns *sy* (*śt*) and *st* may explain why the *īšst* form of the interrogative is more common, but also why the form *īšsy* is also possible.

Allen’s (2000) hypothesis implies that the interrogative *īšst/īšsy* represents a lexicalization of an original cleft structure, such as *What is it(=she) [(that)...]?*. Note that in Old Egyptian *īšst* is particularly common in a cleft construction based on the demonstrative *pw* (see Edel 1964:517 and Section III.2.3.3.2.2), as in (106), which is similar to that of (97b).

Old Egyptian

- (106) *īšst pw ḥm n-n irrw-tn?*
 IPW COP indeed N-DEM do.IPFV.REL.M-2PL.SUF
 ‘What is it that you are doing?’ (Edel 1964:517)⁷⁴

The link between the interrogatives *īḥ* and *īš-st/īš-sy* and the noun (*iḥ*)-*t* ‘thing(s), something, anything, property’ (also ‘products, dishes, meals, sacrifice, ritual, party, celebration’, cf. Vycichl 1984:20) suggests a development ‘thing(s), property, something, anything’ → ‘what?’.

It should be noted that Allen’s (2000) description of the third person dependent pronouns is somewhat different from that of Gardiner (2001:45 [1957]), as well as Loprieno (1996:67). Although at first sight the two analyses appear to contradict each other, they do not differ that radically in substance. The important element in Allen’s description that helps to reconcile them is the hedge “for the most part”. Let us briefly consider Gardiner’s description of the three third person dependent pronouns, *sw*, *sy* (*śt*) and *st*. According to Gardiner, all three can refer to both people and things, while the latter two, *sy* (*śt*) and *st*, are both third person *feminine* singular dependent pronouns. Their interpretation as feminine is supported by the fact that *-t* is a regular nominal feminine marker in Egyptian, as well as in Afro-Asiatic in general. Furthermore, by Late Egyptian, *st* remains the only third person singular feminine dependent pronoun (cf. Loprieno 1996:67). Its frequent use in reference to things and abstract notions will then simply reflect the Old and Middle Egyptian preference for the feminine in such cases, as already mentioned in Sections III.2.3.3.2.1-III.2.3.3.2.2. Note in this respect that feminine (overtly) marked by the affix *t* is commonly used in Afro-Asiatic languages for abstract notions and “socially inactive” entities (Diakonoff

⁷⁴ The masculine agreement on the verb here is controlled by the neuter demonstrative *nn*. According to Gardiner (2001:417, 86 [1957]), the demonstratives of the *n*-series trigger masculine (singular) agreement on “participles and relative forms” but “the resumptive pronoun then used is feminine”.

1965:53 [1988]; cf. also Achab 2005:102-103 for Berber). Thus, it is quite plausible that at a certain point a tendency existed for the two feminine singular dependent pronouns, *sy* (*sī*) and *st*, to be differentiated in the way indicated by Allen (2000), the latter being more frequently used as ‘she’ rather than ‘it’, while the former, on the contrary, as ‘it’ rather than ‘she’. By Late Egyptian, when the masculine singular dependent pronoun mostly took over the ‘it’ function, the two feminine dependent pronouns were hardly functionally differentiated anymore and the form *st* has ousted the form *sy* (*sī*).

In both Old Egyptian (Jean Winand, p.c.) and Middle Egyptian (107), *išst/išsy* triggers the feminine rather than the masculine agreement pattern.

Middle Egyptian

(107) *išsy pw iry-t?*

IPW COP/DEM do.PFV-F

‘What is to be done?’ (Gardiner 2001:407 [1957], Callender 1975:97)

As a rule, the interrogative *išst/išsy* means ‘what?’, so that Gardiner (2001:407 [1957]) explicitly prefers to interpret the only example he has where it could mean ‘who?’ (108) not as ‘Who is he who is there?’ but as ‘What is (the matter with) him who is there?’ (‘since there is no clear evidence that *išst* ever means ‘who?’’).

Middle Egyptian

(108) *išst pw nty im?*

IPW COP/DEM REL(M) there

‘What is (the matter with) him who is there?’ or maybe ‘Who is he who is there?’ (Gardiner 2001:407 [1957]).

The larger context of this example, which comes from *The Story of the Eloquent Peasant*, may fit Gardiner’s interpretation rather well, yet it does not provide any definitive evidence in favour of it either. Typologically, the usage of the ‘what is he?’ as ‘what’s up?’-construction is not implausible (cf. Section I.4.2.3.3.3). For Old Egyptian, Edel (1955:90, 1964:517-518) glosses *išst* as both ‘what?’ and ‘who?’. The example he provides for the meaning ‘who?’ is reproduced in (109).

Old Egyptian

(109) *išst pw t3w (i)dd(i) iri-f m tr?*






IPW COP man say.IPFV.(M) do.PFV.(M)-3SG.M.SUF in time

‘Who is a man who says that he can do it on time? (the original German translation: *Wer ist ein Mann, der sagt, dass er (es) zur (rechten) Zeit schafft?*)’ (Edel 1964:517), or rather, ‘What is a man who says that he can do it on time?’, ‘What kind of man would/can say that he can do it on

time?', or even 'What's up with a man who says/can/would say that he can do it on time?'⁷⁵

However, (109) can also be interpreted as (i) a KIND-question with 'what?'-dominance (cf. Section II.2), i.e. 'What is a man who says that...?' meaning 'What kind of man would/can say that...?', or maybe it is even similar to Gardiner's interpretation of (108) as (ii) a 'what's up?'-construction 'What's up with a man who says/can/would say that...?'. At least, the use of the indefinite article with *Mann* 'man' in Edel's translation very much favours the latter two interpretations (compare also (112)). Hannig (2003:220-221) gives examples with *īšst* meaning both 'what?' and 'who?', but these appear to be just the same examples as those cited by Gardiner (2001 [1957]) and Edel (1964). Other sources I consulted gloss *īšst/īšsy* only as 'what?'. Summing up, it seems to be highly unlikely that *īšst/īšsy* can be used as 'who?' in Old and Middle Egyptian.

2.3.3.2.4 The interrogative *zy* and the like

The interrogative summarized in (85e) is transliterated in a variety of ways. First, there is variation between *z* and *s* and, second, between *y* and *ī*. The variation between *z* and *s* in its transliterations is due to the fact that the original Old Egyptian distinction between *z* and *s* (phonologically, probably /θ/ vs. /s/, cf. Allen 2000:16) disappeared by Middle Egyptian so that only *s* has remained, but because Egyptian orthography was rather conservative, the distinction was often still preserved in writing. As a rule, the hieroglyph used to write the interrogative at issue is  (O34, the so-called "bolt") or  (O35, from a combination of O34 with "walking legs" sign D54). The original phonogram value of both signs is *z*. Therefore, transliterations with *z* reflect an older reading, whereas transliterations with *s* reflect a later merger of *z* with *s*. Note, however, that since this interrogative was only very rarely written with a phonogram  *s* (S29) in Middle Egyptian (Gardiner 2001:407 [1957]), it seems to be more appropriate to transliterate it with *z*. The variation between *y* and *ī* in transliterations of this interrogative appears to be a matter of interpretation, for it does not reflect any difference in hieroglyphic form. In both cases, the writing is either  (double M17) or  (Z4). Since in Middle Egyptian the two writings normally have the same transliteration value *y* (Gardiner 2001:29, 481, 536-537 [1957]), it seems reasonable to follow Gardiner in transliterating this interrogative with *y*.

The interrogative *zy* is common in attributive use as 'which?', 'what (kind of)?', as in (110).

⁷⁵ The initial and the final *ī*'s in the verb *īdī* 'say' are put in brackets because they do not appear to be represented in the hieroglyphic writing that Edel gives for this example, nor are they transliterated in an identical example found in Hannig (2003:220-221).

Middle Egyptian

(110) *zy/s(y) w3t?*

IPW/IPW path

‘which path?’ (Allen 2000:55), ‘[on] what road?’ (Gardiner 2001:407 [1957])

Allen (2000:55) considers *zy* to be “the first noun of a direct genitive”, that is a genitive construction with a genitive marker *n(y)* ‘of’, whereas Gardiner (2001:407 [1957]) suggests that the noun following *zy* is “in apposition to it”. This common attributive usage may account for the final *-y* of *zy*, for there are good chances that it is the same *-y* as the adjectivizing suffix *-y* (cf. Gardiner 2001:61-62 [1957]). Note, however, that *zy* never agrees in gender-number with the noun it modifies (Gardiner 2001:407 [1957]).

Next to its attributive use as ‘which?, what (kind of)?’, *zy* is also claimed to be used independently as ‘who?’ or ‘what?’, though “not very frequent[ly]” in Middle Egyptian (Edel 1955:90, 1964:518, Gardiner 2001:407 [1957]). However, in many cases *zy* could also be interpreted as ‘which one?’, which matches its basic attributive usage somewhat better. This is in fact how Callender (1975) and Allen (2000) prefer to gloss *zy*. They do not give the glosses ‘who?’ and ‘what?’ at all. Nevertheless, I believe that this is an oversimplification because some examples hardly allow for a ‘which one?’ reading. Compare, for instance, examples (111-113).

Old Egyptian

(111) *zi n-n iddy?*

IPW N-DEM say.IPFV.(M)

‘Who is this who is speaking?’ (Edel’s 1964:518 translation is *Was ist das da, was spricht?* (= *wer spricht?*))(112) *zi pw (i)dd(i) iri-f m tr?*

IPW COP/DEM say.IPFV.(M) do.PFV.(M)-3SG.M.SUF in time

‘Which one/Who can/would say that he can do it on time?’ (Edel’s 1964:518 translation is *Wer ist einer, der sagen (kann), dass er es zur (rechten) Zeit schafft?*)

Middle Egyptian

(113) *ntk sy?*

2SG.M.INDEP IPW

‘Who are you?’ (Gardiner 2001:407 [1957])

Also interesting here are examples (114-116).

Middle Egyptian

- (114) *sy pw ntr msy m min?*
 IPW COP/DEM god bear.PFV.M in today
 ‘Who is the god born today?’ (Gardiner 2001:407 [1957]), ‘Which (one) is the god who has been born today?’ (Callender 1975:97), or ‘What god has been born today?’
- (115) *sy pw mi-w p-w ʿ3?*
 IPW COP cat-M M-DEM great
 ‘What is that great cat?’ (Depuydt’s 1999:238 translation ‘Who is that great cat?’)
- (116) *sy ty pw ʿt-y ipt-f?*
 IPW actually COP limb-DU DEM.F.PL-3SG.M.SUF
 ‘What are those two limbs?’ (Gardiner 2001:407 [1957])

In (114), *zy* could still be interpreted attributively, since *pw* when used as a copula/identificational demonstrative “tends to move after the first prosodic unit of the sentence, regardless of its position in the semantic structure, even in cases when this leftward movement breaks the surface entity of a phrase” (Loprieno 1996:104; cf. (100) and (102) above). The structure of (114) would then be ‘What god [*zy ntr*] (is) it [*pw*] (which) has been born today [*msy m min*]?’.

However, examples (115) and (116) can hardly be interpreted this way because of the presence of a demonstrative modifier, (the second) *pw* and *iptf*, respectively.

The interrogative *zy* appears to function as a nominal predicate that can mean both ‘who?’ (111-113) and ‘what?’ (114-116). The interrogative *zy* can be compared to the French interrogative *quel* ‘which [N]?, what (kind of) [N]?’ which is usually used attributively, as in *quel arbre/homme?* ‘which tree/man?, what (kind of) tree/man?’, but also allows for predicative use as either ‘what?’ or ‘who?’ (cf. Section III.3.1.3.1).

2.3.3.3 Late Egyptian

The Late Egyptian interrogative pronominals are summarized in (117). In what follows, I base myself on Černý (1978), Erman (1968 [1933]), Junge (1996), Korostovtsev (1973), and Lesko & Lesko (2002, 2004). Similarly to Old and Middle Egyptian, there is only one interrogative in Late Egyptian that is regularly used attributively, viz. *it*. When other interrogatives are translated as adjectival ‘what (kind of)?, which?’, they are normally used in a genitive construction. The interrogative *ptr* appears to be rare, as it is given only by Lesko & Lesko (2004:159). It is probably best viewed as a remnant of the Middle Egyptian stage. Due to the lack of examples I will not discuss it.

(117) *Late Egyptian interrogative pronominals (based on Černý 1978, Erman 1968 [1933], Junge 1996, Korostovtsev 1973 and Lesko & Lesko 2002, 2004)*

a. <i>nym</i> (<i>nīm</i>)		‘who?’
<i>m</i>		‘who?, what?’
b. <i>ih</i>		‘what?’, maybe rarely ‘who?’
c. <i>it(i)</i>		‘which [N]?’, ‘where?’, ‘which one?, who?’
d. <i>ptr</i>		‘what?, where?’

2.3.3.3.1 The interrogatives *nym* and *m*

The form *m* is mentioned only by Lesko & Lesko (2002:167), who gloss it as ‘who?, what?’, and is probably best viewed as a remnant of the Middle Egyptian stage similar to *ptr*. The interrogative *nym* is reported to mean only ‘who?’. It can also be used to ask about a person’s name (118) (cf. Section II.3.3.1).

Late Egyptian

(118) *nym rn n p3yī it?*

who name of my father

‘What (lit.: ‘who?’) is the name of my father?’ (Erman 1968:376 [1933])⁷⁶

The form *nym* results from a combination of the subject focus particle *in* with the interrogative *m(i)* giving later the Coptic interrogative *nīm* ‘who?’ (Gardiner 2001:406 [1957], Edel 1964:515, Till 1986:102 [1966]; cf. Section III.2.3.3.5). As it appears from the possibility to use *nym* with prepositions, as a direct object (cf. Korostovtsev 1973:82), or with the subject focus particle *in* (Jean Winand, p.c.), by Late Egyptian it has already fused into a monomorphemic word. The medial sign (Z4) *y(i)* in *nym* (*nīm*) is generally believed to render the vowel *i*.

2.3.3.3.2 The interrogative *ih*

As mentioned in Section III.2.3.3.2, the interrogative *ih* is first reported in Middle Egyptian and it becomes common by the Late Egyptian period. As already mentioned in Section III.2.3.3.2.3, according to Gardiner (2001:408 [1957]), the interrogative *ih* goes back to the word (*i*)*h*-*t* ‘thing(s), something, property’ (-*t* is the feminine suffix). In Middle Egyptian, *ih* is only reported to

⁷⁶ Transliteration of this example is mine because Erman gives only the hieroglyphic form.

mean ‘what?’ and mostly in combination with prepositions (Gardiner 2001:408 [1957]), as in (119).

Middle Egyptian

- (119) *hr ih?*
 on IPW
 ‘why?, because of what?’ (Gardiner 2001:408 [1957])

In Late Egyptian, the earlier syntactic restrictions are lifted. However, *ih* is only rarely used as subject (Korostovtsev 1973:80), as in (120).

Late Egyptian

- (120) *iy ih iw(-i) mi n-3 hms-kwi?*
 come.PFV IPW AUX(-1SG.SUF) like N-DEM sit-RES.1SG
 ‘What will happen while I am sitting (here) like this?’ (Korostovtsev 1973:80)⁷⁷

The use of *ih* as object is illustrated in (121).

Late Egyptian

- (121) *iw-i r irt ih r-sn?*
 AUX-1SG.SUF to do.INF IPW to-3PL
 ‘What shall I do to them?’ (Korostovtsev 1973:80)

Most sources on Late Egyptian gloss *ih* only as ‘what?’. However, Černý (1978:34-35) also glosses it as ‘who?’, as in (122).

Late Egyptian

- (122) *n-y sw ih t-3 ip-t?*
 of-ADJ 3SG.M.DEP IPW F.SG-DEM box-F
 ‘To whom does it – the box – belong?’ (Černý 1978:35) or maybe ‘To what does it – the box – belong?’⁷⁸

It should be noted, though, that it is the only such example he provides and since the context is lacking, *ih* in (122) could also be interpreted as ‘what?’, the overall meaning being then ‘To what does it – the box – belong?’, i.e. ‘What does it – the box – make part of?’ or ‘Where does it – the box – belong?’. As appears from example (123), *ih* ‘what?’ can also be used in a KIND-question (cf. Sections II.2.1.3-II.2.1.4).

⁷⁷ The form glossed as resultative in this example is the so-called “old perfect” in traditional Egyptological terminology (cf. Korostovtsev 1991:291-292).

⁷⁸ The pronoun *sw* in this construction refers to the possessed irrespective of the gender-number of the latter.

Late Egyptian

- (123) *ih hr ib-k n-n n '3m.w?*
 IPW to heart-2SG.M.SUF N-DEM of Asian-M.PL
 ‘What are these Asians for you?’ (Erman 1968:374 [1933])

2.3.3.3.3 The interrogative *it(i)*

According to Erman (1968:376 [1933]), *it(i)* is found “only in texts of refined language” (“nur in Texten gewählter Sprache”), which suggests that it is not a frequent interrogative in Late Egyptian. Hoch (1994:43-44) claims that *it(i)* is of Semitic origin. He reconstructs it as **ē-dē* < ***ay-dē*, where the first part is a widespread Semitic interrogative root ‘which (one)?, where?’ and the second part is a demonstrative, similar for instance to the Classical Arabic demonstrative root *dā*. Reinforcement of interrogatives by means of deictic elements (demonstrative or personal pronominals) is quite common in Afro-Asiatic in general and in Semitic in particular (e.g., see Barth 1913:137-150).

In its use and semantics, the interrogative *it(i)*, as in (124), is rather similar to the Old and Middle Egyptian interrogative *zy* (see Section III.2.3.3.2.4).

Late Egyptian

- (124) a. *rr it(i) dmi?*
 beside IPW town
 ‘beside which/what town?’ (Erman 1968:377 [1933], Hoch 1994:44)
- b. *hr t3-y-f it(i) rwi3-t?*
 on F.SG.DEM-ADJ-3SG.M.SUF IPW side-F.SG
 ‘On which side of it [is the town of *Hirba*]?’ (Hoch 1994:44, Erman 1968:377 [1933])

When *it(i)* is used independently, as in (125),⁷⁹ the sources gloss it as ‘who?’.

Late Egyptian

- (125) *it(i) stn-f n hm-k?*
 IPW equal-3SG.M.SUF to majesty-2SG.M.SUF
 ‘Who is like Your Majesty?’ (Erman 1968:377 [1933], Hoch 1994:44, Korostovtsev 1973:83)

However, a ‘which one?’ interpretation cannot be excluded until a larger context can be considered. Remarkably, no source I consulted glosses *it(i)* as ‘what?’ and in this respect it differs from *zy*. Moreover, Lesko & Lesko (2002:52) report the meaning ‘where?’ for *it(i)*, which *zy* lacks.

⁷⁹ Note that this is the only example used by three sources to illustrate the independent use of *it(i)*.

2.3.3.4 Demotic Egyptian

Demotic interrogative pronominals are summarized in (126). Due to practical difficulties I cannot reproduce the Demotic writings of these interrogatives. In what follows, I base myself on du Bourguet (1976), Johnson (2000, 2001), Lexa (1947-1951) and Spiegelberg (1925).

(126) *Demotic interrogative pronominals (based on du Bourguet 1976, Johnson 2001, Lexa 1947-1951 and Spiegelberg 1925)*⁸⁰

- a. *nm, nim(e)* ‘who?’
- b. *ih* ‘what?’

These appear to be the only two (simple) interrogative pronominals attested in Demotic. Syntactically, they behave like nouns. The interrogative *ih* is also often used in a genitive construction with (or sometimes without) the marker *n*, (127-128), where *ih* functions as a modifier, even though syntactically it forms the head of the genitive construction.

Demotic Egyptian

- (127) *ih n rm t3y?*
 IPW of person F.SG.DEM
 ‘What kind of person is she?’ or ‘Who is she?’ (Spiegelberg 1925:19, Lexa 1947-1951:187)
- (128) *ih (n) i' b3?*
 IPW of illness
 ‘What illness?’ (Spiegelberg 1925:19)

Johnson (2000:15) claims that both *ih* and *nm* “meant either ‘who?’ or ‘what?’”, but unfortunately she does not provide any examples of such versatility.

Since examples in the sources are presented out of any context, it is somewhat difficult to say with certainty whether *nim* and *ih* can be used as ‘which one?’ in questions about humans and non-humans respectively. In this respect, consider example (129), where *nim* is used pronominally.

⁸⁰ In Demotic data, some sources use the transliterations *a* and *e* because of their usual Coptic outcomes, ⲁ and ⲉ, (in hieroglyphic writing, these transliterations are rendered as ⲛⲁ (M17-Z7) and ⲛⲉ (D21-Z1), otherwise they are transliterated as *r* and *i(w)*, respectively (Lexa 1947-1951:44, du Bourguet 1976:3-4). Other transliterations that sources may differ on are *'* ~ *i* and *i* ~ *y*, the latter variant in both cases is the same as in transliterations of earlier Egyptian data.

Demotic Egyptian

- (129) *nim tī t3 nt h(r) t3 mhw3(-t) n p3imw?*
 IPW F.SG.DEM F.SG.DEM REL under F.SG.DEM household of PROP
 ‘Which is this (ship) that carries the tribe of Pemou?’ (Lexa 1947-1951:188) or maybe ‘Who is in charge of Pemou’s household/family?’

Lexa (1947-1951:188) translates it as *Quel est ce (navire) qui porte la tribu de Pemou?* ‘Which is this (ship) that carries the tribe of Pemou?’ or ‘Which one (the ship) carries the tribe of Pemou?’. If Lexa’s translation is correct than this example may be interpreted as a forerunner of a usage of *nim* that becomes somewhat more common later in the Coptic period (perhaps under Greek influence). However, since no context is provided other interpretations are also possible. For instance, the preposition *h(r)* can be interpreted not as ‘under [N], carrying [N]’, but as ‘in the charge of [N]’, corresponding then to the Middle Egyptian *hr-^f (n)* ‘in the charge of (lit.: ‘under the hand of’)’ (Gardiner 2001:587, 132 [1957]). Consequently, the translation will be ‘Who is in charge of Pemou’s household/family?’.

2.3.3.5 Coptic Egyptian

Coptic interrogative pronominals are summarized in (130). In brackets, I indicate the dialect for which a given form is attested. An asterisk preceding the name of the dialect means that a given form is considered by Vycichl (1984) to be “aberrant” for this dialect. Sahidic and Bohairic have been the two major dialects in the history of Coptic. In what follows, I base myself on Crum (1962 [1939]), Černý (1976), Eberle & Schulz (2004), Elanskaya (1991), Lambdin (1983), Plumley (1948), Steindorff (1979), Till (1961, 1986 [1966]) and Vycichl (1984).

The interrogative *nim* has its origin in an Old and Middle Egyptian combination of the subject focus particle *in* and the interrogative *m(i)* ‘who?, what?’, which appears to have fused in the meaning ‘who?’ already by the Late Egyptian period (Edel 1964:515, Gardiner 2001:406 [1957], Till 1986:102 [1966]; cf. Sections III.2.3.3.2.1 and III.2.3.3.3.1). The interrogative *aš* is usually assumed to go back to an Ancient Egyptian interrogative *iḥ* ‘what?’ (Vycichl 1984:20, cf. Sections III.2.3.3.2.3 and III.2.3.3.3.2). The interrogative *ou* must be a Coptic innovation because it is not attested in any earlier form of Egyptian. Vycichl (1984:228) suggests that the interrogative *ou*, together with a formally identical singular indefinite article, goes back to the numeral ‘one’, which in Middle Egyptian was written as *w^f* and had the forms *oua* (M) and *ouei* (F) in, for instance, Sahidic Coptic. In Coptic, the numeral is also used as an indefinite, but only about persons as ‘(a certain) one, a certain man/woman, someone’ (Lambdin 1983:64, Till 1986:105-106 [1966]).

(130) *Coptic interrogative pronominals (based on Crum 1962 [1939], Lambdin 1983, Plumley 1948, Till 1961, 1986 [1966] and Vycichl 1984)*⁸¹

- | | | | |
|----|-----|---|---|
| a. | ⲛⲓⲙ | <i>nim</i> (Bohairic, Sahidic) | ‘who?’, ‘which (one)?
(person or thing)’ |
| b. | ⲟⲩ | <i>ou</i> (Sahidic, Bohairic) | ‘what?’, ‘what (kind of) [N]’ |
| | ⲟⲩⲟ | <i>ouo</i> (Akhmimic) | (thing)?’ |
| | ⲟⲩⲱ | <i>ouō</i> (*Sahidic) | |
| | ⲟⲩⲁ | <i>oua</i> (*Oxyrhynchite) | |
| | ⲟ | <i>o</i> (Akhmimic, Assyutic) ⁸² | |
| | ⲱ | <i>ō</i> (Akhmimic, *Sahidic) | |
| | ⲉⲩ | <i>eu</i> (Assyutic, Sahidic) | |
| | ⲟⲩⲛ | <i>oun</i> (Fayyumic, *Sahidic) ⁸³ | |
| c. | ⲁⲩ | <i>aš</i> (Bohairic, Sahidic) | ‘what (kind of) [N]?, which |
| | ⲉⲩ | <i>eš</i> (Assyutic, Fayyumic) | [N]?, ‘which one (thing)?, |
| | ⲉⲩ | <i>ex</i> (Akhmimic) | what? (predicative)’ |
| | ⲁⲩ | <i>ax</i> (Sahidic _p) ⁸⁴ | |

In the remainder of this section, I will examine the semantics and the patterns of use of each of the three Coptic interrogative pronominals in more detail, viz. *ou* in Section III.2.3.3.5.1, *nim* in Section III.2.3.3.5.2, and *aš* in Section III.2.3.3.5.3. In most sources the glosses of all the three interrogatives would be quite similar, ‘who?’, ‘what?’ and, mostly for *nim* and *aš*, also ‘which [N]?, what (kind of) [N]?’. This is in fact rather misleading, especially because the number of examples provided in the sources to illustrate the glosses is usually quite limited. For instance, cf. Crum (1962 [1939]), Černý (1976), Eberle & Schulz (2004), Macomber (n.d./1999?), Plumley (1948), Steindorff (1979), Till (1961, 1986 [1966]) and Vycichl (1984). There is of course some variation here as well

⁸¹ In representing Coptic data, I use the standard transliteration, as can be found in Lambdin (1983:x) or Elanskaya (1991:301). Here are some important conventions: (i) the transliteration of ⲟⲩ is *ou*, phonologically it is presumed to be /u ~ u:/ or /w/, (ii) the transliteration of ⲉⲩ is *eš*, phonologically it is presumed to be /i ~ i:/ or /j/ and sometimes /ej ~ əj/, (iii) the transliteration of ⲉ is *e*, phonologically it is presumably /ε/ or /ə/. The so called “supralinear stroke”, as in ⲛ̄ and ⲩ̄, is preserved in the transliteration, *n̄* and *š̄*. There is some disagreement among specialists whether the stroke renders a schwa before (or sometimes, after) the consonant or whether it marks the syllabic character of the consonant (cf. Lambdin 1983:xiii-xv). The letter ⲭ is transliterated as *ǧ*.

⁸² Assyutic is also known as Subakhmimic or Lycopolitan.

⁸³ According to Jean Winand (p.c.), there are no examples of this form.

⁸⁴ Sahidic_p stands for Vycichl’s (1984:xii) Sahidic “of proverbs” (“dialecte des Proverbes”), or a “protodialect of Sahidic” (“protodialecte du sahidique”), which represents one of the older versions of Sahidic and for which only one text was found.

and some sources provide glosses more sparingly than others. Till (1961, 1986 [1966]) and Lambdin (1983) are perhaps the most meticulous in their choice of glosses. In glossing the Coptic interrogative pronominals in (130), I based myself primarily on the latter three sources, as well as on my analysis of the examples provided in Crum (1962 [1939]).⁸⁵

The evidence for a lack of differentiation between ‘who?’ and ‘what?’ in Coptic is meager. Of the three interrogatives, only *aš* has some chances to eventually prove to be a ‘who?, what?’ interrogative we are looking for.

2.3.3.5.1 The interrogative *ou*

The interrogative *ou* represents the most straightforward case. Generally, it means ‘what?’ and it is typically used as an object (also of a preposition), as in *r ou* ‘do what?’, and somewhat less frequently so as a nominal predicate (cf. Plumley 1948:165), as in (131).

Coptic Egyptian

- (131) a. *ou-ou* *pe?*
 INDF.ART.M.SG-IPW M.SG.this
 ‘What is it?’ (Lambdin 1983:19, Till 1986:102 [1966])
- b. *hen-ou* *ne?*
 INDF.ART.PL-IPW PL.this
 ‘What are these?’ (Lambdin 1983:19, Till 1986:102 [1966])

It is not particularly common either in modifying use as the syntactic head of a genitive construction marked by *n* ‘of’ (cf. Lambdin 1983:62), as in *ou m̄-mine n* [N] ‘what sort of [N]?, what kind of [N]? (lit.: ‘what of kind of [N]?)’. In the *ou n*-N construction, *ou* means ‘what (kind of) [N]?’ rather ‘which [N]?’. Moreover, it seems that N never denotes a human being here. Finally, the interrogative *ou* also appears to be used in KIND-questions, as in (132).

Coptic Egyptian

- (132) *ntk ou-ou?*
 2SG.M INDF.ART.M.SG-what
 ‘What are you?’ (Lambdin 1983:19, Till 1986:102 [1966])

⁸⁵ It should be noted that Crum (1962 [1939]) is a very specific kind of dictionary. As a rule, instead of concrete examples it just gives a reference to the original source. Fortunately, most of such references relate to various texts of the Bible, which are relatively easy to check. But even when Crum (1962 [1939]) does provide concrete examples, these mostly come without translations, with all the resulting inconveniences.

Example (132) may be compared to a similar Late Egyptian example (123) in Section III.2.3.3.3.2. In all probability, it is due to examples like (132) that *ou* is glossed as ‘who?’ in many sources. Note, however, that both Lambdin (1983:19) and (Till 1986:102 [1966]) translate (132) as ‘What are you?’ and not ‘Who are you?’. It is also quite revealing that the sources provide no examples of *ou* being used in the meaning ‘who?’ as a subject or object.

2.3.3.5.2 The interrogative *nim*

The basic meaning of the interrogative *nim* is ‘who?’, which it has maintained since the Late Egyptian stage. It can have this meaning when used as an object (also of a preposition) (133) and nominal predicate (134).

Coptic Egyptian

- (133) *p-šēre* *ñ-nim?*
 DEF.ART.M.SG-son of-IPW
 ‘Whose son?’ (Lambdin 1983:19, Till 1986:102 [1966])

- (134) *nim* *pe* *pei-rōme?*
 IPW M.SG.COP M.SG.this-man
 ‘Who is this man?’ (Lambdin 1983:19)

I have not found examples of *nim* being used as a subject probably because a cleft construction is normally used in questions about subjects of verbs (135), where *nim* is best analyzed as a nominal predicate.⁸⁶

Coptic Egyptian

- (135) *nim* *pe-ntaf-tsabe-tēutñ...?*
 IPW M.SG.this-PRF₂.3SG.M-teach-2PL
 ‘[O generation of vipers,] who hath warned you (lit.: ‘who (is it) that have warned you’) [to flee from the wrath to come]?’ (Matthew 3:7 in Wells 2000-2006)⁸⁷

Furthermore, the interrogative *nim* can be used as the syntactic head of a genitive construction marked by *n* ‘of’, as in (136), where it can be glossed either as ‘which [N]?’ or ‘what (kind of) [N]?’, depending on the larger context.

⁸⁶ According to Jean Winand (p.c.), there are some examples where *nim* is used as a subject.

⁸⁷ In glossing *pe* here as a demonstrative ‘this(M.SG)’, I follow Eberle & Schulz (2004:70). This gloss is probably due to the fact that it is also possible to insert a second *pe* before the first one, the second *pe* being the copula. Recall, however, that in the Middle Egyptian cleft construction the same masculine demonstrative *pw* is usually analyzed as a copula (cf. Section III.2.3.3.2.1).

Coptic Egyptian

(136) a. *nīm n-rōme*

IPW of-man

‘Which man?’ or ‘What (kind of) man?’ (Lambdin 1983:62, Till 1986:102 [1966])

b. *nīm m-prohoimion*

IPW of-preface

‘What (kind of) preface?’ or ‘Which preface?’ (Crum 1962:225 [1939])⁸⁸

Remarkably, *nīm* in the *nīm n*-N construction can be used both with human nouns (136a) and non-human nouns (136b), although the latter use is less common according to Plumley (1948:164). Finally, as a nominal predicate, *nīm* can be used selectively, both in questions about humans and non-humans, corresponding to English ‘which one?’. Thus, Crum (1962:225 [1939]) reports that there are some Sahidic and Bohairic versions of Matthew 23:17, ‘Ye fools and blind: for whether is greater, the gold, or the altar that sanctifieth the gold?’, which use *nīm* instead of the regular *aš*, as in (144) below.

No example was found in the sources where a predicatively used *nīm* could be unambiguously glossed as ‘what?’. The only exception may be its use in NAME-questions (Lambdin 1983:19, Till 1986:102 [1966]), as in (137), but this use does not constitute a lack of differentiation (cf. Section II.3). Recall that a similar usage is attested already in Late Egyptian (cf. example (118) in Section III.2.3.3.3.1).

Coptic Egyptian

(137) *nīm pe pe-k-ran?*

IPW M.SG.COP M.SG-2SG.M.POSS-name

‘What is your name?’ (Lambdin 1983:19, Till 1986:102 [1966])

2.3.3.5.3 The interrogative *aš*

The interrogative *aš* is particularly common in two positions. First, it is used as the syntactic head of a genitive construction marked by *n* ‘of’, (139-142), where it can mean both ‘which [N]?’ and ‘what (kind of) [N]?’ depending on the larger context, although the latter meaning seems to be more common.

⁸⁸ The translation is mine.

Coptic Egyptian

- (139) *aš n̄-hi pe-tetna-kot-f?*
 IPW of-house(M) M.SG.this-FUT.2PL-build-3SG.M
 ‘What (kind of) house are you going to build?’ (Till 1986:103 [1966])
- (140) *n̄-aš n̄-he akč-ine m̄mo-f?*
 in-IPW of-manner PRF.2SG.M-find OBJ-3SG.M
 ‘How did you find him?’ (Lambdin 1983:19, Till 1986:102 [1966])
- (141) *ou-aš n̄-he pe pei-maein?*
 INDF.ART.M.SG-IPW of-manner M.SG.COP M.SG.this-sign
 ‘Of what sort is this sign?’ (Lambdin 1983:62) or ‘What (kind of) sign is this?’
- (142) *ou-aš m̄-mine pe pei-rōme?*
 INDF.ART.M.SG-IPW of-kind M.SG.COP M.SG.this-man
 ‘Of what sort is this man?’ (Lambdin 1983:62) or ‘What (kind of) man is this?’

As far as I can judge from the data provided in the sources, the *aš n̄-N* construction is hardly ever used with human nouns, (143) being the only possible example I encountered.

Coptic Egyptian

- (143) *aš/nim gar n̄-sarks ie nim...?*
 IPW because of-flesh then all
 ‘For who (is there) of all flesh (i.e., mortal men), [that hath heard the voice of the living God speaking out of the midst of the fire, as we have, and lived]?’ (Deuteronomy 5:26; the reference is from Crum 1962:22 [1939], the text is available at <http://www.remenkimi.com>)⁸⁹

However, even here much depends on the exact interpretation of the word *sarks* ‘flesh’ (compare also example (174) in Section III.2.3.4.1.3.4, which is a Biblical Hebrew version of the same example). Note that it is also possible to use *nim* instead of *aš* in (143).

Second, *aš* is common in predicative use where it can usually be glossed as ‘which one?’ or ‘what?’, as in (144) and (145).

⁸⁹ The two *nim*’s, ‘who?’ and ‘all, every’ have different etymologies (Vycichl 1984:142).

Coptic Egyptian

- (144) *aš gar pe pnoč?*
 IPW because M.SG.COP great
 ‘[Ye fools and blind:] for whether is greater, [the gold, or the altar that sanctifieth the gold]?’ (Matthew 23:17 in Eberle & Schulz 2004:15, Till 1986:102 [1966], Wells 2000-2006)
- (145) *aš ne?*
 IPW PL.this
 ‘(And the one of them [...] answering said unto him, Art thou only a stranger in Jerusalem, and hast not known the things which are come to pass therein these days? And he said unto them), What things? (lit.: ‘what (are) these?’)’ (Luke 24:18-19 in Till 1986:102 [1966])

Despite the fact that many sources also gloss *aš* as ‘who?’, the predicative *aš* appears to be used with humans as extremely rarely as *aš* in the *aš n-N* construction. Examples (146) and (147) are in fact the only two examples I was able to find in the sources.

Coptic Egyptian

- (146) *aš te te-k-shime?*
 IPW F.SG.COP F.SG-2SG.M.POSS-wife
 ‘Who is your wife?’ (Steindorff 1979:47)
- (147) *aš pe p-rōme?*
 IPW M.SG.COP DEF.ART.M.SG-man
 ‘Who is the man?’ (Eberle & Schulz 2004:24)

If the translations of (146) and (147) in the sources are correct, Coptic *aš* would be functionally very much like the Old and Middle Egyptian interrogative *zy* (Section III.2.3.3.2.4; e.g., (147) can be compared to (113)). However, since no context is provided, *aš* in (146) and (147) may also be selective, viz. ‘Which one is your wife?’ and ‘Which one is the man?’. It is also possible that (146) and (147) are KIND-questions, similar to (132) with *ou* above, viz. ‘What is your wife?’ and ‘What is the man?’ respectively. I have not found any examples of *aš* being used on its own as an object.

2.3.4 Semitic languages

The Semitic branch of Afro-Asiatic used to comprise two major sub-branches, Western and Eastern. However, the Eastern languages, Akkadian (with its later Assyrian & Babylonian varieties) and Paleosyrian (Eblaite & Amorite varieties), ceased to be spoken already in the antiquity. Depending on the classification of Arabic, the Western languages are subdivided in two branches either as

Northwest Semitic vs. South Semitic (including Arabic) or, more recently, as Central Semitic (= Northwest Semitic + Arabic) vs. South Semitic (see Faber 1997). As I am not in position to judge the classifications proposed, I will arbitrarily follow the second classification.

I will begin this section by a general overview of the Semitic interrogative pronominals in Section III.2.3.4.1. In Section III.2.3.4.2, I will examine Semitic languages where one interrogative pronominal seems to be indiscriminately used as ‘who?’ and ‘what?’. I will present data from some Modern Arabic varieties spoken in Northwest Africa, Syria and Southern Mesopotamia (Section III.2.3.4.2.1), Soqotri (West Semitic, South, Eastern; Yemen; Section III.2.3.4.2.2), and the “Canaano-Akkadian” mixed language (an extinct mixed East-West Semitic idiom; ca. 1350 BC; Section III.2.3.4.2.3). Of the three cases, that of Southern Mesopotamian Arabic is clearly the most solid one. The other two cases remain somewhat questionable for the time being, because the data available is rather fragmentary. Canaano-Akkadian is extinct with only a limited record left, and Soqotri, although still very much alive, is seriously underdescribed. Furthermore, the mixed nature of the Canaano-Akkadian language presents some additional difficulties for an univocal interpretation of the available data.

2.3.4.1 Semitic interrogative pronominals: an overview

2.3.4.1.1 Forms, meanings, history

Thanks to the long history of records of various Semitic languages,⁹⁰ we have data on many languages already extinct by now, as well as on previous stages of several modern Semitic languages. The interrogative pronominals of various Semitic languages are summarized in Table 1. As far as possible, I also give the forms of the attributive interrogative ‘which [N]?, what (kind of)?’.⁹¹ The latter is clearly related to the widespread Semitic root *ʔay ‘where?’ (Lipiński 1997:328). In various Afro-Asiatic languages this ‘where?, which (one)?’ interrogative appears to have developed into the interrogative pronominal ‘who?’, particularly in Cushitic and South Omotic, or ‘what?’, especially in North Omotic (cf. Dolgopolsky 1991:12, Bender 2000:208-209, and Table 2 in Section III.2.3.5.1).

⁹⁰ For instance, the first written records of the East Semitic languages, Akkadian and Paleo-Syrian, are dated to ca. 2300 BC (Buccellati 1997:69, Gordon 1997:101).

⁹¹ Admittedly, in some languages this interrogative may behave rather like a nominal. For instance, in Classical Arabic it occurs in the so-called “construct state” and the noun it modifies follows in the genitive (Fischer 1997:197), as in *ʔayy-u rajul-in* ‘which man?’ (which-NOM.SG.M.CS man(M)-GEN.SG.INDF). What is important, however, is that *ʔayyu* is prototypically used together with another nominal (the genitive), for which it functions as a modifier.

Table 1. Some of the Semitic interrogative pronominals (mostly based on the papers in Hetzron 1997 and on Lipiński 1997:328-329, other sources include Abu-Absi 1995, Barth 1913, Bittner 1916, Brockelmann 1908, 1913, Müller 1905, Titov 1991, and personal communication from Charles Häberl, David Kummerow, Marie-Claude Simeone-Senelle)⁹²

Language	‘who?’	‘what?’	‘which (one)?’
East Semitic			
Old Akkadian	<i>man</i>	<i>min</i>	
Old Babylonian	<i>mannu</i> (NOM)	<i>mi:nu</i> (NOM)	<i>ayy-</i>
Eblaite Paleosyrian	<i>mi:</i>		
Amorite Paleosyrian	<i>manna</i>	<i>ma:</i>	<i>ayy-</i>
West Semitic, Central			
Old Canaanite	<i>miya</i>	<i>manna</i>	
Ugaritic	<i>my</i>	<i>mh/mn</i>	<i>ʔay</i>
Biblical Hebrew	<i>mi(y)</i>	<i>ma(h) ~ mo(h)</i>	<i>ʔay(-)</i>
Aramaic	<i>man(nu)/mannu:</i>	<i>ma:(h)</i>	<i>ʔay-</i>
Syriac Aramaic	<i>man</i>	<i>ma:</i>	<i>ʔayna:</i>
Neo-Mandaic Aramaic	<i>man</i>	<i>mu ~ mo</i>	<i>hem</i>
Ṭuroyo Neo-Aramaic	<i>man</i>	<i>min</i>	
Classical Arabic	<i>man</i>	<i>ma: ~ mah</i>	<i>ʔayy-</i>
Cairene Arabic	<i>mi:n</i>	<i>ʔe:h</i>	<i>ʔa:y/ʔanhu</i>
Damascus Arabic	<i>mi:n</i>	<i>šu:(we)/ʔe:š</i>	<i>ʔayy/ʔanu</i>
Jewish Baghdadi Arabic	<i>manī</i>	<i>ašku:n/aš e:š</i>	<i>hayyi</i>
Negev Bedouin Arabic	<i>min</i>	<i>e:š/i:š wiš</i>	<i>ya:t</i>
Yemenite Arabic	<i>man/min</i>	<i>ma:(di)/weyš we:ššu</i>	<i>ʔayyan</i>
Moroccan Arabic	<i>(a)škun</i>	<i>š(nu)/aš</i>	<i>ina</i>
Chadian Arabic	<i>ya:tu</i>	<i>šunu</i>	<i>we:nu</i>
West Semitic, South			
Mehri	<i>mo:n/éy</i>	<i>hé:śən</i>	
Shehri	<i>mu(n)</i>	<i>iné ~ íne</i>	
Bathari	<i>ma:n</i>	<i>híne</i>	

⁹² At times, the interrogative pronominals in the table and in the examples may be marked somewhat inconsistently, especially vowel length (e.g., *a ~ a:*, *u ~ u:*) and sometimes quality (e.g., *ä ~ a ~ ε*, *ə ~ ɪ*), which is largely due to different notations used in various sources. I have tried to unify them.

Geez	<i>mannu(:)</i> (M.NOM)	<i>mint</i> (M.NOM)/ <i>mi(:)</i>	<i>ay</i>
Tigre	<i>man</i>	<i>mi(:) ~ mə</i>	<i>ʔay-</i>
Tigrinya	<i>män</i>	<i>min</i>	<i>ʔintaway</i> , <i>ʔäyyänay</i> (persons); <i>ʔintay</i> , <i>mintay</i> (things)
Amharic	<i>man</i>	<i>min(din)</i>	<i>yät-ḥṇaw</i> (M.SG, things or persons); <i>man-ḥṇaw</i> (persons),
Dahalik ⁹³	<i>man</i>	<i>mi</i>	<i>iy-</i> (things or persons); <i>man</i> (persons)
Harari	<i>ma:n</i>	<i>min</i>	<i>a:y</i>
Silte	<i>ma:</i>	<i>min</i>	<i>ayta:y</i>
Outer South Ethiopian	<i>m^wa(n)</i>	<i>mən ~ mər/ mə-qar</i>	<i>mən ~ mər</i>

In general, Semitic interrogative pronominals look routinely the same, the most noticeable exception being some secondary variation in the ‘what?’ interrogatives of modern Arabic dialects (cf. Section III.2.3.4.2.1) and Modern South Arabian languages. The interrogative pronominals meaning ‘who?’ typically show up either as *ma(n) ~ ma:(n)* or, in Northwest Semitic languages and Eblaite Paleosyrian, as *miy(a) > mi(:)* (Lipiński 1997:328), while the interrogative pronominals meaning ‘what?’ typically show up as *mah(a) > ma:* (Lipiński 1997:328), or, in Akkadian and Ethiopian South Semitic languages, as *mi(n) ~ mi:(n)* (Modern Arabic forms do not belong here, see below for more details). The loss of final *n*, or on the contrary, its increment as an original part of some “reinforcing” deictic (cf. Section III.2.3.4.1.2), contraction of *y* or *h* with or without some compensatory lengthening, and other similar processes may give an impression that the forms of the interrogative pronominals swap their functions somewhat too freely across the Semitic languages: the same forms may mean ‘who?’ in some languages but ‘what?’ in some others. A conclusion that might then suggest itself would be that, originally, Semitic languages did not make any distinction between ‘who?’ and ‘what?’. Consider, for instance, the following note from Brockelmann (1908:326).

Während die semit[ischen] Sprachen sonst beim Nomen und Pron[omen]

⁹³ Dahalik (West Semitic, South) is a modern Semitic language spoken on the Dahlak Archipelago off coast of Eritrea. It was discovered only in 1996 (see Simeone-Senelle 2005).

nur zwei ideelle Rangklassen, Genera, unterscheiden, ist bei den substantivischen Fragewörter ein Gegensatz zwischen Person und Sache ausgebildet; doch fällt diese Unterscheidung erst in die Periode des einzelsprachlichen Lebens, daher dieselben Elemente in den verschiedenen Sprachen z[um] T[eil] in entgegengesetzter Bedeutung auftreten.

The “flip-flopping” tendency is particularly impressive if, as Brockelmann, or Barth (1913) normally do, only the first syllables of the interrogative pronominals, *ma(:-)* and *mi(:-)*, are compared, the second consonant *n* being disregarded as not pertaining to the etymological root. However, even if fuller forms are compared, a good deal of apparent “flip-flops” could still be found relatively easily, especially if we take a larger Afro-Asiatic perspective.⁹⁴ Alternatively, we may choose to follow Lipiński (1997:328), who splits the interrogative pronominals as *m-* and the rest, *-an* and *-iy(a) > -i:* for ‘who?’ and *-in* and *-ah(a) > -a:* for ‘what?’,⁹⁵ which will leave us with only one root *m* ‘who?, what?’.

As has been suggested above, many of the “flip-flops” are rather likely to be only apparent and due to secondary developments, such as the loss of final *n*, etc. At the same time, I believe that the structural analyses for the etymological developments of the Semitic interrogative pronominals proposed by Brockelmann, Barth, and particularly Lipiński, are valid in many cases as well. The following facts are probably most relevant here. First, note the omnipresence of the initial *m-* in the interrogative pronominals and the existence of various deictic roots based on the same consonants as the second parts of the interrogative pronominals (cf. Barth 1913:72-77, 89-103; Brockelmann 1908:316-324; Militarev & Stolbova, n.d., the latter also for other branches of

⁹⁴ For instance, we have, on the one hand, Eblaite Paleosyrian *mi:* ‘who?’ (East Semitic; Syria; ca. 2300-2250 BC; Lipiński 1997:328), Ugaritic *my* ‘who?’ (West Semitic, Central, Northwestern; Syria; ca. 1400-1190 BC; Pardee 1997:134), Old Canaanite of El-Amarna letters *miya* ‘who?’ (West Semitic, Central, Northwestern; data for the period ca. 1350 BC; Segert 1997:178), and Biblical Hebrew *mi(y)* ‘who?’ (West Semitic, Central, Northwestern; Israel). On the other hand, we have Tigre *mi(:) ~ mə* ‘what?’ (West Semitic, South, Western, Ethiopian; Eritrea; Barth 1913:141; Militarev & Stolbova, n.d.; Raz 1997:448), an infrequent Geez form *mi(:)* ‘what?’ (West Semitic, South, Western, Ethiopian; Ethiopia; ca. 350-1000 AD; Barth 1913:141; Militarev & Stolbova, n.d.), Burji *miya* ‘what?’ (Afro-Asiatic, East Cushitic, Highland; Ethiopia; Hudson 1976a:259), Proto-Chadic *mi* ‘what?’ (Frajzyngier 1985:64), Malian Tuareg *mi* ‘who?’ and ‘where? (what?)’ (Afro-Asiatic, Southern Berber; Mali; see Section III.2.3.2.2.3), and Ait Ndir *mi* ‘who?, what?’ (Afro-Asiatic, Northern Berber; Morocco; see Section III.2.3.2.2.3).

⁹⁵ Although Lipiński does not pronounce on the possible origins of these post-*m* elements, he does call them “morphemes”.

Afro-Asiatic).⁹⁶ Second, there is a clear tendency in Semitic languages for conventionalizing the structure [interrogative + deictic], which is probably best analyzed as clefting with the deictic functioning as a copula (cf. Section III.2.3.4.1.2). Moreover, this tendency has a good number of parallels in other branches of Afro-Asiatic (cf. Section III.2.3.2 for Berber and Frajzyngier 1985:64-66 for Chadic). Third, Berber languages, a relatively closely related branch of Afro-Asiatic, provide a good example of a system of interrogative pronominals based on the structure [a general interrogative pronominal root *m* + a deictic] (cf. Section III.2.3.2). The way the system is organized in Zenaga of Idab Lahsan (Afro-Asiatic, West Berber; Mauritania; cf. Section III.2.3.2.2.5) may be particularly interesting for a reconstruction of the Proto-Semitic system. Thus, it is not unlikely that Proto-Semitic possessed one interrogative pronominal meaning only ‘who?’ and at least one interrogative that could mean ‘who?’ or ‘what?’ depending on the context, and when needed, disambiguated through the choice of a suitable deictic/copula. The first interrogative was probably **ma(:)n* ‘who?’, as can be deduced from the wide presence of this form with this meaning in both East and West Semitic. The other interrogative pronominals were most likely **mi* ‘who?, what?’ and **ma* ‘what?’. They were typically followed by a “reinforcing” element and must have become specialized in ‘who?’ vs. ‘what?’ meanings rather early, **mi-y(a) > mi(:)* ‘who?’, **mi-n(a) > min* ‘what?’, and **ma-h(a) > ma(:)* ‘what?’. These three forms suggest that on the Proto-Semitic level **ma* was already specialized as ‘what?’, but **mi* was ambiguous between ‘who?’ and ‘what?’. It seems unlikely that similarly to **mi-n(a) > min* ‘what?’, **ma(:)n* ‘who?’ would come from **ma-n(a)*. The reason is that in this case there would be two ‘who?, what?’ interrogatives, **mi* and **ma*, which in combination with the same element **n(a)* would produce the opposite meanings, viz. ‘what?’ and ‘who?’.⁹⁷

⁹⁶ Semitic deictic roots are numerous and tend to form compounds, with their vocalic components often expressing case, “state”, gender-number distinctions. Establishing their original meanings and paths of semantic evolution would be a rather onerous enterprise falling outside the scope of the present study. Yet, it seems reasonable to assume that the Proto-Semitic deictic system was at least no less complex than that of the modern Berber languages (Section III.2.3.2.2.1). Probably, it distinguished minimally between various degrees of distance, grammatical gender, number, anaphoric reference, perhaps also humanness, definiteness and modifying/pronominal use. Furthermore, there must have been some neutral forms, unmarked for gender, number and/or other categories.

⁹⁷ Instead, the interrogative pronominal **ma(:)n* ‘who?’ may be cognate to an East Cushitic word for ‘man, person’, e.g. *manni* ‘person (NOM)’ in Libido (Afro-Asiatic, East Cushitic, Highland; Ethiopia; Joachim Crass, p.c.) or *maa-* ‘man’ in Daasanach (Afro-Asiatic, East Cushitic, Western Omo-Tana; Ethiopia; Sasse 1976:207). More similar forms meaning ‘man, people’ from other branches of Afro-Asiatic can be found in Militarev & Stolbova (n.d.). Other

Besides the interrogative pronominals mentioned above, Proto-Semitic also had an interrogative root ‘where?’ **ʔay* (or the like), which could be used pronominally in the meaning ‘which one?’ and/or attributively as ‘which [N]?, what (kind of) [N]?’. Apparently, in its pronominal use it was regularly followed by **ya*, giving **ʔay-ya*, similarly to **mi-y(a)* ‘who?’.⁹⁸ This link between ‘which one?’ and ‘who?’ is not surprising (cf. Section I.2.3). In this respect, consider also Cushitic and South Omotic where *ʔayy-* (or the like) typically means ‘who?’ (cf. Dolgopolsky 1991:12, Bender 2000:208-209, and Table 2 in Section III.2.3.5.1)

By way of conclusion, let us examine a few forms of interrogative pronominals whose meaning in a given language is the opposite of the meaning this form normally has in the family. Such cases look like “flip-flops”, but most likely they are not. For instance, consider the Syriac Late Aramaic (West Semitic, Central, Northwestern; Syria; 200-1200 AD) form *ma:n(a:)* ‘what?’, which looks very much like a widespread Semitic interrogative *ma(:)n* ‘who?’. However, according to Barth (1913:140), this Syriac Late Aramaic interrogative goes back to a combination of *ma:* ‘what?’ and a demonstrative *dēn(a:)* (this combination is also attested in Jewish Literary Late Aramaic, also known as Targumic). Somewhat similarly, we have in Ugaritic (West Semitic, Central, Northwest; Syria; ca. 1400-1190 BC) *mn* ‘what?’, in Old Canaanite *ma-an-na* (transliteration) /manna/ (usual transcription; Lipiński 1997:328), and in older Biblical Hebrew a rare interrogative *mān* ‘what?’ (Gesenius & Kautzsch 1962:119; Macuch 1969:254). Lipiński (1997:328) explains the Ugaritic and the Old Canaanite forms as contractions from **mahna*. The latter form he compares to the (Minaic/Minean) Epigraphic South Arabian (West Semitic, South, Western; first millennium BC – first half of the first millennium AD) interrogative *mhn* ‘what?’, which he suggests to vocalize as **mahna*. In some languages, the gemination of *-n* in the *man*-like ‘who?’ interrogatives may be due to the fusion with a demonstrative, e.g. *ha(:)*, or personal pronominal, e.g. 3SG.M *hu(:)*. For instance, in Aramaic *man* ‘who?’ was often used with the enclitic 3SG.M pronoun *hu:*, which resulted in the form *mannu:* ‘who?’ (Kaufman 1997:122).

The Modern Arabic form *mi(:)n* ‘who?’ is formally (almost) identical to the Old Akkadian and common Ethiopian Semitic ‘what?’ interrogatives. Yet, the Modern Arabic form must represent a development of the Old Arabic *man*. The

hypotheses may be possible as well, but this does not really matter for the moment.

⁹⁸ The element **ya* is likely to be of a deictic origin. For similar demonstratives in Semitic, see Barth (1913:89-91, 115-116, 129-130), for Egyptian Loprieno (1996:68), for Chadic Frajzyngier (1985:64-66). Alternatively, it may be cognate to the Proto-Cushitic auxiliary **y ~ *a* ‘be, say’ (cf. Cohen *et al.* 2002:234). However, since copula verbs often develop from deictics in Afro-Asiatic, the latter verb may ultimately have a deictic origin as well.

Modern Arabic form *mi(:)n* probably results from the so-called *ʔima:la* “/a/-raising”, a change from /a/ to /e ~ i ~ ie/ due to the “umlauting influence” of /i/ (Kaye & Rosenhouse 1997:279). Compare, in this respect, the Jewish Baghdadi Arabic form *mani* ‘who?’ (Kaye & Rosenhouse 1997:290) or the Christian Neo-Aramaic of Urmi *ma:niy* ‘who?’ and Kerend Neo-Aramaic *máni:* ‘who?’ (West Semitic, Central, Northwestern; Iran; Jastrow 1997:355). The final *i*, which is not present in the Classical Arabic and Aramaic interrogative *man* ‘who?’, may have various sources: a genitive case ending,⁹⁹ a fused demonstrative, personal or relative pronominal, copula,¹⁰⁰ to name just a few. However, it is more likely that in Modern Arabic varieties this *i* is of epenthetic origin. For instance, in Modern Standard Arabic, *i* is the “default juncture [i.e., epenthetic] vowel” used to prevent “non-canonical clusters of consonants arising between words in juncture” (Badawi et al. 2004:10). In the case of *man* ‘who?’, there is one quite frequent context where such “non-canonical cluster” would occur, namely when *man* is followed by a singular masculine definite relative pronominal (*a*)*lladi:*, as in (148).¹⁰¹

Modern Standard Arabic (West Semitic, Central; Badawi et al. 2004:693)

- (148) *man-i lladi: sa-yu-ša:b-u [bi-rtifa:ʔi ɖagʔi l-dami]?*
 IPW-EP DEF.REL.SG.M FUT-3SG.M.SBJ-afflict.IPRF.PASS-3SG.M.SBJ
 ‘Who is the one who will be afflicted [with a rise in blood pressure]?’

According to Xavier Luffin (p.c.), the first vowel of the relative pronominal (*a*)*lladi:* is regularly dropped when it is preceded by another word, which without the insertion of an epenthetic *i* would create a forbidden consonant cluster. The same is true for the definite article *al-* and definite phrases are not infrequent following *man*. Together with the strong preference of Arabic for syllables with onsets, this would also regularly result in a forbidden three consonant cluster **ma[n_I-CV...]*.

⁹⁹ The Classical Arabic form *man* does not seem to inflect for case and the Modern Arabic varieties usually lack case marking altogether. However, of the three Classical Arabic case endings, the nominative *-u*, the accusative *-a*, and the genitive *-i*, the varieties which do show some traces of case marking normally preserve only *-i* (cf. Kaye & Rosenhouse 1997:284, Lipiński 1997:264).

¹⁰⁰ For instance, in Aramaic the 3SG.M bound pronominal is *-eh ~ -hi:* (Kaufman 1997:122). In Neo-Mandaic, one of the survived descendants of Aramaic, the same bound pronominal has the form *-i* (Häberl 2006:173). Neo-Mandaic also has a 3SG.M “enclitic pronominal copula” *-(y)ye*, which is “most commonly used with question words” (Häberl 2006:267-268).

¹⁰¹ I am grateful to Xavier Luffin, Marie-Claude Simeone-Senelle and Martine Vanhove for suggesting the gloss for the form *man-i* in this example.

2.3.4.1.2 Deictic “reinforcement”

Reinforcement of interrogatives by means of deictic elements is quite common in Semitic, as in (149-150) with demonstratives and (151-152) with personal pronominals. In this respect, see, for instance, Barth (1913:137-150) or Brockelmann (1913:196).

Classical & Modern Standard Arabic¹⁰²

- (149) *wa-man da: ya-fu:t-u* *l-mawta?*
and-IPW DEM 3SG.M.SBJ-escape.IPRF-3SG.M.SBJ DEF.ART-death.F.ACC.SG
‘Who can escape death?’ (Brockelmann 1913:196)
- (150) *fa-ma: ha:da: ya-ru:ʿ-u-ka?*
so-IPW DEM 3SG.M.SBJ-scare.IPRF-3SG.M.SBJ-2SG.M.OBJ
‘So, what makes you afraid?’ (Brockelmann 1913:196)
- (151) *man huwa l-ʿafḍalu bayna-hum?*
who 3SG.M DEF.ART-best.M.SG among-3PL.M
‘Who is the best among them?’ (Badawi *et al.* 2004:693)
- (152) a. *ma: huwa l-sababu?*
IPW 3SG.M DEF.ART-cause(M).NOM.SG
‘What is the cause?’ (Badawi *et al.* 2004:694)
- b. *ma: hiya l-ṣu:ratu* [*llati: tatašakkalu huna:*]?
IPW 3SG.F DEF.ART-image.F.NOM.SG
‘What is the cause [which is being formed here]?’ (Badawi *et al.* 2004:694)

Brockelmann (1913:196) remarks that in the Semitic languages “question words can acquire a particular emphasis through demonstratives”, as in (150) and (151), while the use of a relative clause represents another means of “asking a question with more insistence”, as in (153).¹⁰³ The relative pronoun is often itself of a demonstrative origin, as *elli:* in (153) (cf. Lipiński 1997:326).

Egyptian Arabic

- (153) *mi:n elli: ḥadoh?*
IPW REL.SG.M take.PRF.3SG.M.SBJ.3SG.M.OBJ
‘Who took it?’ (Brockelmann 1913:196)

It is also important to mention that different interrogatives tend to select different

¹⁰² Modern Standard Arabic is “the modern equivalent of Classical Arabic, the language defined by medieval grammarians and in principle still the norm for the Arab Academies in making their decisions about [Modern Standard Arabic]” (Badawi *et al.* 2004:2).

¹⁰³ Brockelmann (1913:196): “Die Fragewörter können durch Demonstrativa besonderen Nachdruck erhalten”, “die Frage eindringlicher zu gestalten”.

deictics. Thus, in Modern Standard Arabic (Badawi *et al.* 2004:691-695), *ma:* ‘what?’ appears to be very common with a deictic root *da:*,¹⁰⁴ so that they are even written in one word, while somewhat less common with personal pronominals, whereas *man* ‘who?’ appears to be rare with *da:*, while quite common with the personal pronominals.

The use of a “reinforcing” deictic is probably best analyzed as clefting. Otherwise, the deictic should be interpreted as a focus marker with no influence on the original syntactic relations. Under the cleft analysis, the deictic itself may be analyzed in a number of ways: as a relative pronoun, as a subject of the main clause in a cleft construction of which the interrogative is the predicate (e.g., $[[it]_{\text{SBJ}} [(is) \textit{who}]_{\text{PRED}}]_{\text{MAIN}} [(that) P]_{\text{SUBORDINATE}}?$), or as a copula. The latter analysis may be preferable because the development from demonstratives to copulas is a common phenomenon in Afro-Asiatic (cf. Section III.2.3.3.2.2 for Egyptian, as well as Frajzyngier 1985:66 for Chadic, who also gives more references). However, what matters most here is the tendency to conventionalize a cleft construction of the structure [interrogative + deictic], which is quite common in other branches of Afro-Asiatic as well.

2.3.4.1.3 When ‘who?’ is *not* ‘what?’ and ‘what?’ is *not* ‘who?’

In the present section I will briefly discuss some uses of interrogative pronominals in Semitic languages that in the sources are often presented as cases of lack of differentiation between ‘who?’ and ‘what?’, and I will argue that this is not the case. The most common case appears to be represented by the use of ‘who?’ in the case of a non-prototypical combination of values of the kind [thing + identification (+ proper name)], viz. in NAME-questions (cf. Section II.3). The use of ‘what?’ in the case of a non-prototypical combination of values of the kind [person + classification (+ common noun)] (or KIND-question; cf. Section II.2) is taken for a lack of differentiation between ‘who?’ and ‘what?’ less frequently, apparently because it is also typical for the native languages of most authors, English, French, and German, although not Russian, as can be seen from the discussion of Soqotri example (43). Furthermore, sources on Biblical Hebrew also typically treat the ‘what’s up?’-construction (cf. Section I.4.2.3.3.3) as a lack of differentiation between ‘who?’ and ‘what?’, as well as the use of *mi(y)* ‘who?’ with “things denoting persons”, as David Kummerow (p.c.) puts it, or in Joüon’s (1965 [1923]:446) words, “pour les choses [...] avec une idée latente de personne” (“for things [...] with a latent idea of a person”).

¹⁰⁴ The “core deictic element” *da:* is regularly “supplemented by other deictic elements”, as in *ha:da:* ‘this.M.SG’, *ha:dihi:* ‘this.F.SG’, *da:lika* ‘that.M.SG’, etc. (Badawi *et al.* 2004:47).

2.3.4.1.3.1 NAME-questions

In many Semitic languages we find ‘who?’ being used in the case of a non-prototypical combination of values of the kind [thing + identification (+ proper name)] (cf. Section II.3), viz. in questions about (personal) proper names, as illustrated in (154-160). Modern Hebrew and a few other modern Semitic languages that use ‘what?’ here seem to form an exception rather than the rule (161-162). As can be seen from (155a) and (155b), in Biblical Hebrew both ‘who?’ and ‘what?’ were possible in questions about personal proper names. The same situation is attested in some Modern Arabic varieties, (163).

Old Babylonian (East Semitic; Iraq; ca. 2000-1500 BC)

- (154) *mann-um šum-ka?*
 who-NOM name-2SG.M
 ‘What is your name?’ (Izre’el & Cohen 2004:111)

Biblical Hebrew (West Semitic, Central, Northwest; Israel; ca. the first millennium BC)

- (155) a. *mi(y) šmε-χ?*
 who name-2SG.M
 ‘What is your name?’ (Judges 13:17 via Brockelmann 1913:195, David Kummerow, p.c.)
 b. *ma-ššmε-χ?*
 what-name-2SG.M
 ‘What is your name?’ (Genesis 32:27 via Brockelmann 1913:195)

Mandaic Late Aramaic (West Semitic, Central, Northwest; Iraq & Iran; 200-1200 AD; Brockelmann 1913:195)

- (156) *mannu: šum-a:k?*
 who.3SG.M name-2SG.M
 ‘What is your name?’

Mehri (West Semitic, South, Eastern; Yemen & Oman)

- (157) *hamm-ek mo:n?*
 name-2SG.M who
 ‘What is your name?’ (Brockelmann 1913:195; see also Simeone-Senelle 1997:414)

Amharic (West Semitic, South, Western, South Ethiopian; Ethiopia)

- (158) *sem-wo man?*
 name-2SG.HON who
 ‘What is your name?’ (Brockelmann 1913:195)

Tigre (West Semitic, South, Western, North Ethiopian; Eritrea)

- (159) *sem-ek man?*
 name-2SG.M who
 ‘What is your name?’ (Brockelmann 1913:195)

Geez (West Semitic, South, Western, North Ethiopian; Ethiopia; ca. 350-1000 AD; Brockelmann 1913:195)

- (160) *mannu: sem-eka?*
 who.NOM name-2SG.M
 ‘What is your name?’

Modern Standard Arabic (West Semitic, Central)

- (161) *ma: ism-u-ka?*
 what name-NOM.SG-2SG.M
 ‘What is your name?’

Modern Hebrew (West Semitic, Central, Northwest; Israel)

- (162) *ma šim-χa?*
 what name-2SG.M
 ‘What is your name?’ (Shlomo Izre’el, p.c.)

Chadian Arabic (West Semitic, Central; Chad)

- (163) a. *ʔusm-ak ya:tu?*
 name-2SG.M who
 ‘What is your name?’ (Abu-Absi 1995:36)
 b. *šunu ʔusm-ak?*
 what name-2SG.M
 ‘What is your name?’ (Abu-Absi 1995:35)

2.3.4.1.3.2 KIND-questions

In rather many Semitic languages, we find ‘what?’ being used in the case of a non-prototypical combination of values of the kind [person + classification (+ common noun)] (or KIND-question; cf. Section II.2), at least in some contexts (164-171).

Classical & Modern Standard Arabic (West Semitic, Central)

- (164) *ma: ha:ʔula:ʔi n-nisa:ʔu?*
 what this.PL DEF.ART-women.NOM
 ‘What are these women?’ (Brockelmann 1913:195)
 (165) *fa-ma: tazawwaj-ta bikr-a-n ʔaw*
 so-what marry.PRF-2SG.M virgin-ACC.SG-INDEF or

ṭayyib-a-n?

deflowered-ACC.SG-INDEF

‘What (woman) have you married, a virgin or an already deflowered one?’
(Brockelmann 1913:195)

- (166) *ma: ʔantuma: [wa-min ʔayna jiʔtuma:]?*

what 2DU.M

‘What sort of people are you both (lit.: ‘What are you both?’) [, and from where did you come] (Fischer 1997:213)

Modern Hebrew (West Semitic, Central, Northwest; Israel)

- (167) *ma aḫi-ḫa, napaḫ o naḡar?*

what brother-2SG.M smith or carpenter

‘What is your brother, smith or carpenter?’ (Shlomo Izre’el, p.c.)

Neo-Mandaic Aramaic (West Semitic, Central, Northwest; Iran)

- (168) *mo-yye ahh-ak, qana:ya: ya: neja:ra:?*

what-3SG.M brother-2SG.M smith or carpenter

‘What is your brother, smith or carpenter?’ (Charles Häberl, p.c.)

Amharic (West Semitic, South, Western, South Ethiopian; Ethiopia)

- (169) *John mindin nä-w?*

PROP what COP-3SG.M

‘What is John?’ (Seyoum Mulugeta, p.c.)

Shehri (West Semitic, South, Eastern; Oman)

- (170) *iné tit-k?*

what wife-2SG.M

‘[And he asked him,] What (kind of person) is your wife? (‘Was ist deine Frau (für eine)?’) [He said to him, ‘She is worth nothing, she has damaged my clothes.’]’ (Bittner 1917:106-107)

- (171) *emé-k b-íné zhoñt?*

mother-2SG.M with-what come.PRF.3SG.F.SUBJ

‘[And he asked him,] What has your mother given birth to? (‘Was hat deine Mutter (zur Welt) gebracht?’) [The guy said to him, ‘She has given birth to a girl...’]’ (Bittner 1917:74-75)

The use of ‘what?’ in KIND-questions seems to be somewhat less widespread than the use of ‘who?’ in NAME-questions. In Modern Standard Arabic and Modern Arabic varieties, the acceptability of ‘what?’ in KIND-questions seems to vary depending on the region. A separate study would be necessary here to determine the exact distribution. The Hebrew language may provide an example of a diachronic change in the acceptability of ‘what?’ in KIND-questions. Thus, whereas this is possible in Modern Hebrew (Shlomo Izre’el, p.c.), (167), it seems

to be impossible in Biblical Hebrew (David Kummerow, p.c.).

2.3.4.1.3.3 The ‘what’s up?’-construction

Some sources on Biblical Hebrew providing examples (172-173) treat them as cases of lack of differentiation between ‘who?’ and ‘what?’ (Macuch 1969:254-255) or KIND-questions (involving ‘what?’-dominance) (Joüon 1965 [1923]:446).¹⁰⁵ However, I believe that they are best analyzed as instances of the ‘what’s up?’-construction (cf. Section I.4.2.3.3.3). This interpretation fits somewhat better the larger contexts of these examples and the general lack of instances of the use of ‘what?’ in KIND-questions in Biblical Hebrew texts. Furthermore, several translations of the same sentences in other languages are clearly based on a ‘what’s up?’ interpretation, as in (172b-e). The fact that examples of ‘what’s up?’-construction are extremely rare in Biblical Hebrew texts¹⁰⁶ may be an indication of its somewhat idiosyncratic or very colloquial nature and/or that it is a very expressive, stylistically marked way of asking such a question, which probably should have been avoided in a biblical text. Compare the similar rarity of ‘what’s up?’-construction in Old and Middle Egyptian texts (Afro-Asiatic, Egyptian; cf. example (108) in Section III.2.3.3.2.3).

Biblical Hebrew

(172) a. *mə(h) hə-šivri(y)m hə-ʔelə(h)?*

what DEF.ART-Hebrew.PL.M DEF.ART-DEM.PL.M

‘[Now the Philistines gathered together all their armies to Aphek: and the Israelites pitched by a fountain which is in Jezreel. And the lords of the Philistines passed on by hundreds, and by thousands: but David and his men passed on in the rereward with Achish. Then said the princes of the Philistines,] What do these Hebrews (here)? (lit.: ‘What (are) these Hebrews?’ → ‘What’s up with these Hebrews?’ or ‘Why are these Hebrews (here)?’) [And Achish said unto the princes of the Philistines, Is not this David, the servant of Saul the king of Israel, which hath been with me these days, or these years, and I have found no fault in him since he fell unto me unto this day?]]’ (1 Samuel 29:3 via Joüon 1965 [1923]:446, transliteration mine); Joüon’s own translation is *Qu’est-ce que ces Hébreux?* ‘What are these Hebrews?’

¹⁰⁵ Macuch (1969:254): “ein ursprüngliches, altertümliches מַה [i.e., ‘what?'] mit Bezug auf Personen”. Joüon (1965 [1923]:446): *ma(h)* ‘what?’ “peut s’employer en parlant d’une personne pour demander *ce* quelle est”.

¹⁰⁶ Thus, Macuch (1969:254) writes that “ein ursprüngliches, altertümliches מַה [i.e., ‘what?'] mit Bezug auf Personen” is attested “an einer einzigen Stelle, Nu 22:9”.

German (via <http://scripturetext.com>)

- b. *Was sollen diese Hebräer?*
 what should.PRES.3PL these Hebrews

Afrikaans (via <http://scripturetext.com>)

- c. *Wat soek hierdie Hebreërs?*
 what seek these Hebrews

Czech (*Bible kralická* version via <http://scripturetext.com>)

- d. *K čemu jsou Židé tyto?*
 for what are Hebrews these

Czech (*Český ekumenický překlad* version via <http://scripturetext.com>)

- e. *Co s těmito Hebreji?*
 what with these Hebrews

Biblical Hebrew

- (173) *mɔ(h) hɔ-ʔānɔʃi(y)m hɔ-ʔele(h) ʕimmɔχə?*
 what DEF.ART-man.PL DEF.ART-DEM.M.PL with.2SG.M

‘[And the elders of Moab and the elders of Midian departed with the rewards of divination in their hand; and they came unto Balaam, and spake unto him the words of Balak. And he said unto them, Lodge here this night, and I will bring you word again, as the Lord shall speak unto me: and the princes of Moab abode with Balaam. And God came unto Balaam, and said,] What men are these with thee? [And Balaam said unto God, Balak the son of Zippor, king of Moab, hath sent unto me, saying, Behold, there is a people come out of Egypt, which covereth the face of the earth: come now, curse me them; peradventure I shall be able to overcome them, and drive them out.]’ (Numbers 22:9 via Macuch 1969:254-255, transliteration mine)

The overall context, especially Balaam’s answer, suggests that the meaning of (173) is, in fact, very similar to that of (172), that is something like ‘What’s up with these men with you?’, ‘What do these men with you want?’, ‘Why are these men here with you?’.

2.3.4.1.3.4 “Things denoting persons”

The use of Biblical Hebrew *mi(y)* ‘who?’ in examples like (174-175) and especially (176-178) is usually presented in the sources as a lack of differentiation between ‘who?’ and ‘what?’.

Biblical Hebrew

- (174) *mi(y) kəl bəśər?*
 who all flesh
 ‘[For] who (is there of) all flesh, [that hath heard the voice of the living God speaking out of the midst of the fire, as we have, and lived]?’ (Deuteronomy 5:26; the reference is due to David Kummerow, p.c.)
- (175) *mi(y) be(y)θ-i(y)?*
 who house.CS-1SG
 ‘[Then went king David in, and sat before the Lord, and he said, Who (am) I, O Lord God? and] what (is) my house [, that thou hast brought me hitherto]?’ (2 Samuel 7:18; the reference is due to Joüon 1965 [1923]:446 and David Kummerow, p.c.)
- (176) *mi(y) ḥayyay mi(y) mišpaḥah ʔv-i(y) bə-yisrʔel?*
 who life.PL.1SG who family.CS father-1SG in-PROP
 ‘[And David said unto Saul, Who am I? and] what (is) my life, (or) my father's family in Israel [, that I should be son in law to the king?].’ (1 Samuel 18:18; the reference is due to David Kummerow, p.c.)
- (177) *mi(y) lə-χɔ kəl ham-maḥānə(h) haz-ze(h)?*
 who to-2SG.M all DEF.ART-camp DEF.ART-DEM.M.SG
 ‘[And he said,] What (meanest) thou (by) all this drove [which I met? And he said, These are to find grace in the sight of my lord]’ (Genesis 33:8; the reference is due to Macuch 1969:254 and Joüon 1965 [1923]:446)
- (178) *mi(y) pešaʔ yaʔāqov? ḥā-lo? šomro(w)n*
 who transgression.CS PROP PQ-NEG PROP
u(w)-mi(y) bəm-o(w)θ yəhu(w)ḏə(h)?
 and-who high.place.CS-PL.F PROP
 ‘[For the transgression of Jacob is all this, and for the sins of the house of Israel.] What (is) the transgression of Jacob? (is it) not Samaria? and what are the high places of Judah? [are they not Jerusalem?].’ (Micah 1:5; the reference is due to Joüon 1965 [1923]:446 and David Kummerow, p.c.)

However, Joüon (1965 [1923]:446) describes this use of *mi(y)* ‘who?’ as “pour les choses [...] avec une idée latente de personne” (“for things [...] with a latent idea of a person”). Similarly, David Kummerow (p.c.) has suggested that *mi(y)* ‘who?’ is used here with “things denoting persons”. I believe that these examples should not be interpreted as questions involving some non-prototypical combination of values either. In the examples at issue *mi(y)* ‘who?’ is not used in any special way. Rather, the nouns it occurs with are used metonymically to stand for a group of people the objects these nouns denote contain, belong to, are produced by, etc. A particular inclination to this kind of metonymy appears to be

just a general feature of the Biblical Hebrew lexicon.¹⁰⁷ The idea that the use of *mi(y)* ‘who?’ in (174-178) represents a lack of differentiation between ‘who?’ and ‘what?’ is largely due to the misleading translations, which use the primary “thing”-meanings instead of reflecting the intended metonymic extensions.

Let us consider these examples in more detail. In (174), the word *bəṣar* (Strong’s 1890 number 01320), derived from a root meaning ‘to be fresh’ (Strong’s 1890 number 01319), is usually glossed as ‘flesh’, but it also has a wide range of metonymic extensions, such as ‘body’, ‘person’, ‘genitals’, ‘kin’, ‘mankind’ (see Strong’s 1890 number 01320). For instance, one of the glosses that Gesenius (1846) gives for *bəṣar* is ‘*all living creatures [...]* especially *all men*, the whole human race [...] Very often used as opposed to God and his power [...], with the notion of weakness and frailty’. Thus, (174) just means ‘Who is there of all (mortal) men [...]?’ (compare in this respect example (143) in Section III.2.3.3.5.3, which is a Coptic Egyptian version of the same sentence). The gloss of the subject root *be(y)θ-* in (175) is ‘house (in the greatest variation of applications, especially *family*, etc.) [*italics added*]’ (Strong’s 1890 number 01004); Gesenius (1846) also gives ‘persons living together in a house, *family*’. Similarly, the word glossed as ‘life’ (Strong’s 1890 number 02416) in (176), derived from a root meaning ‘to live, to nourish up’ (Strong’s 1890 number 02421), is also used to mean ‘relatives’, ‘community’, and ‘sustenance, maintenance’. That is, in (176) the question is about the group of people which “sustains”, “nourishes”, “gives life to” David. In (177), the word glossed as ‘camp (of travellers or troops)’ (Strong’s 1890 number 04264) can also stand metonymically for those who encamp, company, body of people.

Example (178) is somewhat less obvious, at first sight. In all probability, this is due to the fact that the words *pešaʿ* and *bəmo(w)θ* are metonymically extended in a rather idiosyncratic way here. For instance, neither Strong’s nor Gesenius’s lexicons mention the metonymic extensions intended in this example. The word *pešaʿ* (Strong’s 1890 number 06588) is regularly glossed as ‘transgression, revolt (national, moral or religious)’; it is derived from the root ‘to break away (from just authority), trespass, revolt, transgress’ (Strong’s 1890 number 06586). The word *bəmo(w)θ* (Strong’s 1890 number 01116) is regularly glossed as ‘high places’; it is derived from a root ‘to be high’. Consider, however, the following comment on this verse:

Jacob - The sons of Jacob, the ten tribes. What - Or, who is the spring, and cause of that overflowing transgression? Of Jacob - The kingdom of the ten tribes, the head of which was Samaria, where the kings had their residence,

¹⁰⁷ This kind of metonymy is not alien to English either, e.g. *Downing Street said in a statement that there had been a “misunderstanding”* or *The White Rose won the battle at St. Albans in 1455 AD*.

where they worshiped idols, and set an example to the rest of the Israelitish kingdom. And what - Or, who is the cause of the high places, and the idolatry there practised? Jerusalem - Which was the chief city of that kingdom, and had the same influence over Judah, as Samaria had on the ten tribes.

Wesley (1754-1765)

Thus, the word *yaśāqov* ‘Jacob’, the name of the father of the twelve patriarchs of the tribes of Israel, stands for the whole people of Israel and the first question asked is in fact ‘Who are Jacob’s descendants (which group) committing a transgression, a sin?’. And the answer suggested is that these are the people of Samaria. The word *bomo(w)θ* ‘high places’ were also used in Biblical Hebrew to refer to high places (mountains, hills) used as places for worship, especially to idols (Gesenius 1846), and in (178) it metonymically refers to the people practising idolatry on the high places. Therefore, the second question asks ‘Who are Jacob’s descendants (which group) in the tribe of Judah practising idolatry on high places?’. And the answer suggested is that these are the people of Jerusalem.

Old Babylonian (East Semitic; Iraq; ca. 2000-1500 BC) is another Semitic language for which an example quite similar to the Biblical Hebrew examples (174-178) just discussed can be cited, (179).

Old Babylonian (East Semitic; Iraq; ca. 2000-1500 BC)

- (179) *abu:bu* *ša* *taqabba:[ninni]* *mannu* *šu:?*
 flood(M).NOM REL say.IPFV.2SG.M[to.me] who.NOM 3SG.M
 ‘The flood which you mention to me, what is it?’ (Izre’el & Cohen 2004:111), or ‘[They broke the cosmic barrier!] The flood which you mentioned, whose is it? [The gods commanded total destruction! Enlil did an evil deed on the people!]’ (The Schøyen Collection MS 5108, <http://www.nb.no/baser/schoyen/4/4.3/432.html#5108>)

The second translation somewhat better matches the overall context. Example (179) comes from the Babylonian *Atra-Hasis* story of a great flood, which according to the commentary accompanying The Schøyen Collection MS 5108, (<http://www.nb.no/baser/schoyen/4/4.3/432.html#5108>) goes as follows. After the gods created the human race to take over hard agricultural work, the humans multiplied and made such noise that the chief god, Enlil, could not sleep. Accordingly he tried to reduce their numbers, first by plague, then by famine. In each case the god Ea (or Enki), who was mainly responsible for creating the human race, frustrated the plan. Enlil then got all the gods to swear to co-operate in exterminating the whole human race by a huge flood. The flood eventually failed because Ea got his favourite, Ziusudra (or Atra-Hasis), to build an ark and so save the human race and the animals. Thus, the question in (179) seems to ask who is behind the flood.

2.3.4.2 Semitic languages with ‘who?, what?’ interrogatives

2.3.4.2.1 Modern Arabic varieties

According to Brockelmann (1908:327), in various Modern Arabic varieties of Northwest Africa and Syria the interrogatives *ʔa:š* and *šu:*, respectively, which both basically mean ‘what?’ and derive from the phrase *ʔayy-u šayʔ-in* ‘which/what(-NOM.SG.M.CS) thing((M)-GEN.SG.INDF)?’, can also be used as ‘who?’ (“tritt es sogar für ‘wer’ ein”). Unfortunately, Brockelmann does not give any examples of this use. However, we should probably not worry too much about that, since in general Brockelmann seems to be rather careful in distinguishing questions involving non-prototypical combinations of values from a lack of differentiation (cf., for instance, 1913:195). A further development of the phrase *ʔayy-u šayʔ-in* can be found in Moroccan and Tunisian Arabic dialects: *ʔa:šku:n* (Brockelmann 1908:327) ~ *(a)škun* (Kaye & Rosenhouse 1997:290) ‘what?’ < **ʔa:š ya-ku:n* ‘what 3sg.m-be.IPRF’ (Brockelmann 1908:327). Interestingly, while *ʔa:šku:n* ~ *(a)škun* means ‘who?’ in these dialects, *ʔa:š* ~ *aš* means ‘what?’ (Kaye & Rosenhouse 1997:290) or ‘what?, who?’ (Brockelmann 1908:327). Even more remarkable is the fact that in Jewish Baghdadi Arabic the same form *ašku:n* means ‘what?’ and not ‘who?’ (Kaye & Rosenhouse 1997:290).

The development of *ʔayy-u šayʔ-in* > *ʔa:š* ~ *aš* *šu:* from ‘what?’ to ‘what?, who?’ might be accounted for by a particular kind of polysemy typical of the word *šayʔ* ‘thing’ in Arabic.¹⁰⁸ Thus, Brockelmann (1908:329) mentions that already in Old Arabic *šayʔ* ‘thing’ was used as ‘something’ but also as ‘somebody’ and that this is also very common in modern dialects (“So wird schon im Altar[abischen] *šayʔ* ‘Sache’ als ‘etwas’ aber auch als ‘jemand’ gebraucht [...] und so ist es in allen neueren Dialekten ganz gewöhnlich”).¹⁰⁹ This explanation could be applicable to the Northwest African ‘who?, what?’ forms just as well.

Another Modern Arabic variety where a single form can be used both as ‘who?’ and ‘what?’ is spoken in Southern Mesopotamia,¹¹⁰ i.e. in Southern Iraq and the neighbouring Khuzestan region in Iran. In Southern Iraq, this feature is

¹⁰⁸ In Northwest Africa, the development of *ʔayy-u šayʔ-in* > *ʔa:š* ~ *aš* from ‘what?’ to ‘what?, who?’ may also be due to a Berber substrate. Recall that in Berber languages ‘who?, what?’ interrogatives are quite common (see Section III.2.3.2).

¹⁰⁹ Cross-linguistically, Arabic is not exceptional in its use of a word like ‘thing’ as an indefinite pro-word or rather a placeholder both for things *and* humans. Thus, in French *machin* from *machine* ‘machine’ can be used as placeholders both for things and humans. In Flemish, *dinge(s)*, derived from *ding* ‘thing’, is used somewhat similarly.

¹¹⁰ I am grateful to Bruce Ingham for drawing my attention to this fact and for suggesting the references.

attested as far north as the town of Hilla on the Euphrates (Ingham 1982:26). Unlike in the Modern Arabic varieties with a similar phenomenon presented above, in Southern Mesopotamian Arabic it is not the original ‘what?’ but the ‘who?’ interrogative which is involved. According to Ingham (1973:549, 1982:83), Southern Mesopotamian Arabic dialects use an enclitic element *-man*,¹¹¹ corresponding to Classical Arabic *man* ‘who?’, as ‘who?’ or ‘which one? (person or thing)’ with verbs, (180-182), and ‘who?’ or ‘what?’ with prepositions, (183-185), with nouns it means ‘[N] of what?’ or ‘[N] of who?, whose?’.

Southern Mesopotamian Arabic

- (180) *ʔa:xið-man?*
I.shall.take-IPW
‘Which shall I take?’ or ‘Who shall I marry?’ (Ingham 1973:549)
- (181) *tri:d tištiri:-man?*
you.want buy-IPW
‘Which do you want to buy?’ (Ingham 1973:549)
- (182) *šifit-man?*
you.saw-IPW
‘Who did you see?’ or ‘Which did you see?’ (Ingham 1973:549, 1982:26)
- (183) *ʔale:-man iyaw?*
for-IPW they.came
‘Who did they come for?’ or ‘What did they come for?’ (Ingham 1973:549)
- (184) *mim-man imsawwa:ya?*
from-IPW he.is.made
‘What is it made of?’ (Ingham 1973:549)
- (185) *mim-man?*
from-IPW
‘from whom?’ or ‘from which?’ (Ingham 1982:84)

The meaning of *-man* is determined by the nature of the verb and by the context. The translations of examples like (182) and (185) provided by Ingham make me suspect, however, that the semantic division between *-man* encliticized to verbs and to prepositions stated above is not that strict, and in fact, in both cases *-man* could mean ‘who?’, ‘which one? (person or thing)’ or ‘what?’, depending on the context.

Besides the enclitic *-man*, Southern Mesopotamian Arabic dialects have

¹¹¹ Rarely, the form is *-min*, as in Rufai‘ Euphrates Bedouin variety (Ingham 1982:84-85).

several other interrogative pronominals.¹¹² Thus, there is an interrogative root *šin-* ‘what?’ and a somewhat less frequent enclitic (or sometimes, proclitic) element *-š-* ‘what?’ (Ingham 1982:83), both deriving from the same phrase *ʔayy-u šay²-in* ‘which/what thing?’, already mentioned in the beginning of this section. The form *šin-* must be followed by an enclitic third person singular pronoun, masculine *-haw* ~ *-hu* or feminine *-hay* ~ *-hi*, with *h* sometimes falling out (Ingham 1973:550, 1982:87). Furthermore, there is a pair of interrogatives *ya(:)-haw* ~ *ya(:)-hu* ‘who (masculine)?’ and *ya(:)-hay* ~ *ya(:)-hi* ‘who (feminine)?’, which can also be used pronominally as ‘which one? (person and sometimes thing)’ (Ingham 1973:547, 550, 1982:87). The original meaning of the latter interrogative is ‘which one? (person or thing)’, as in Baghdadi Arabic where *ya-hu* is opposed to *man* ‘who?’ (cf. Ingham 1973:549). The root *ya(:)-* clearly corresponds to the Classical Arabic *ayy-* ‘which [N]?, what (kind of) [N]?’.

Why has *man* been extended from ‘who?’ to ‘which one? (person or thing), what?’ in Southern Mesopotamian Arabic? I suppose that this change has been triggered by another development, that of *ya(:)-hu/ya(:)-hi* from ‘which one? (person or thing)’ to both ‘which one? (person or thing)’ and ‘who?’, which in itself is quite natural (cf. Section I.2.3), but results in an overlap in the meanings of *ya(:)-hu/ya(:)-hi* and *man* ‘who?’. Apparently, this overlap has triggered an extension of *man* to ‘which one? (person or thing)’ and later even ‘what?’, subsequently tending to oust the clitic *-š-* ‘what?’. In its turn, *ya(:)-hu/ya(:)-hi* has gone even further in its specialization for questions on humans. In other words, in Southern Mesopotamian Arabic a kind of swap in functions between the two interrogatives has occurred, with the third interrogative, *-š-* ‘what?’ falling into disuse. The perseverance of the non-clitic *šin-* ‘what?’ seems to indicate that the process of turning *man* into an enclitic has also played a role in this evolution.

2.3.4.2.2 Soqotri

Soqotri is one of the so-called Modern South Arabian languages (West Semitic, South, Eastern) spoken in Yemen on the island of Soqatra. Despite the fact that the first publications on Soqotri date as far back as Müller (1902, 1905, 1907), this language still remains poorly described. The most comprehensive grammatical overviews to date are probably Johnstone’s (1975) and Simeone-Senelle’s (1997) sketches of the Modern South Arabian languages. The data on Soqotri presented in this section come from the texts published by Müller (1902, 1905) and Naumkin & Porkhomovsky (1981), as well as from Leslau’s (1938) Soqotri lexicon, which is largely based on Müller’s texts. I am also grateful to

¹¹² There seems to be a great deal of both inter- and intra-dialectal variation here. “[A] particular item may be used in one area but infrequently, whereas in the other area it is the most common term” (Ingham 1982:86).

Marie-Claude Simeone-Senelle and Vladimir Agafonov for discussion and references.

Soqotri interrogative pronominals with their basic meanings are summarized in (186).

(186) *Soqotri interrogative pronominals (based on Leslau 1938, Müller 1902, 1905, Naumkin & Porkhomovsky 1981, Simeone-Senelle 1997)*

‘who?’ *mon, man* (after prepositions: *m^hon, mehon*)

‘what?’ *ínem, (ʔ)íném, ním, ʔíním* (after prepositions: *(ʔ)ín(e)hem, ínʔihem*)

As can be seen, there is a good deal of formal diversity here, which may be due to (morpho)phonologically conditioned variation, as well as to dialectal variation. Soqotri has “a very rich dialectology for which sociological and geographical parameters are relevant” (Simeone-Senelle 1997:379). I will use the first forms, *mon* and *ínem*, for the purpose of reference further on. The *h* in the post-prepositional forms of the interrogatives is most likely the so-called “parasite *h*”, “non-etymological and non-morphological *h* [...] related to the particular evolution of the long vowels and to the rules of stress” (cf. Simeone-Senelle 1997:384). According to an etymology in Leslau (1938:67), the interrogative *ínem* ‘what?’ derives from an older **ayna-m(a)*, where the first element is probably an interrogative ‘which (one)?’ (or ‘where?’), while the latter seems to be either an interrogative pronominal ‘what?’ similar to the Arabic *ma(:) ~ mah*, which is then used as a relative pronominal, or an indefinite pronominal formative, as the Old Akkadian *-ma* or Arabic *-ma:* (cf. Lipiński 1997:330).¹¹³ If this etymology is correct, then *ínem* ‘what?’ has originally been structurally similar to a rare Classical Arabic interrogative *ʔay-ma* ‘what?’ (cf. Barth 1913:129).

Besides the basic meanings presented in (186), the interrogative pronominals are sometimes attributed absolutely opposite secondary meanings by the sources. Thus, Leslau (1938), basing himself on examples in Müller’s texts, defines *ínem* as ‘what?’ (187, 188) and *mon* as both ‘who?’ (189) and ‘what?’ (190, 191).

Soqotri

(187) *ʔínem néfaḥ di-ho di bébe?*

what job of-1SG of father

‘What is my father’s job?’ (Müller 1905:51)

¹¹³ Of course, the indefinite pronominal formative may itself go back to the interrogative ‘what?’. Another plausible origin would be a conjunction, ‘and’ or ‘or’, which in Semitic languages frequently have the forms *wa ~ ma* and *ʔaw ~ wa ~ ʔam* respectively (cf. Lipiński 1997:470-473).

- (188) *ste* *ʔinem?*
 you.are.eating what
 ‘What are you eating?’ (Müller 1907:17)
- (189) *mon í:?* [*ʔeméroh ho Fá:timeh férehim di níbi*]
 who 2SG.F
 ‘Who are you? [She said, ‘I am Fatima, the Prophet’s daughter]’ (Müller 1905:105)
- (190) a. *mon di beʔári-š?*
 IPW REL.SG happen.at.night.IPRF.3SG.M.SBJ-2SG.F.OBJ
 [*wu-lehé be-qáʔer yeʔómer*] *ál boʔer*
 NEG happen.at.night.PRF.3SG.M.SBJ
 ‘[When one comes into a house in the morning, he says,] Good morning! (lit.: ‘What was with you.F.SG at night?’) [And those in the house say,] ‘Good morning!’ (lit.: ‘Nothing happened at night’)’ (Müller 1905:365 lines 27-28; see also Leslau 1938:92)
- b. *mon éy-kin?*
 IPW at-2PL.OBJ
 [*yeʔómer elhé il be-qáʔer*] *bíši díah kál díye*
 there.is.not bad all good
 ‘[When one comes (again) when the sun is high, after he (already) came to the people in the morning and wished them a good morning, he says,] How are you? (lit.: ‘What is at you.PL?’) [And those in the house say,] ‘There is nothing bad, everything is good!’ (Müller 1905:366 lines 7-9; see also Leslau 1938:292)
- c. *mon di qehébi?*
 IPW REL.SG happen.at.noon.IPRF.3SG.M.SBJ
 [*we-yeʔómer elehé il be-qáʔer*] *ál qohób*
 NEG happen.at.noon.PRF.3SG.M.SBJ
 ‘[And when he comes to them after lunch again, he says,] What was at noon? [And those in the house say,] ‘Everything was okay! (lit.: ‘Nothing happened at noon’)’ (Müller 1905:366 lines 14-15; see also Leslau 1938:369)
- (191) *mon dé séher néfog* *ʔe-š hər?*
 IPW this.M.SG moon go.out.PRF.3SG.M.SBJ at-2SG.F today
 [*ʔeméroh séher di ʔésereh*]
 ‘[A question to a pregnant woman:] In which month are you now? (lit.: ‘(For) today, what is this moon (that) has (already) appeared for you?’) [She said, ‘The tenth month’.]’ (Müller 1905:103 lines 6-7)

In turn, Naumkin & Porkhomovsky (1981), while glossing *mon* in Russian only

as *kto* ‘who?’ in their texts, explicitly qualify *ínem* as an “interrogative word *čto* ‘what?’ (in some cases *kto* ‘who?’)” (1981:34). I found only two examples, (192) and (193), where *ínem* is glossed as *kto* ‘who?’ in their texts.

Soqotri

- (192) *Ṭawi:se, Ṭawi:se, wi-nim da di*
 PROP PROP and-IPW this.M.SG REL.SG
ye-wiyór-hen?
 3SG.M.SBJ-follow.IPRF-1PL.OBJ
 ‘Tawise, Tawise, what/who (is) this who is following us?’ (Naumkin & Porkhomovsky 1981:31)
- (193) *ʔínim éke min ʔembóriyo?*
 IPW remain from children
 ‘[If a Soqotri says that his sister has passed away, he will be asked:]
 What/who is left of children?’ (Naumkin & Porkhomovsky 1981:86)

However, not all examples where the interrogative pronominals are assigned a meaning opposite to the usual one seem to stand up to a closer examination. Let us first consider examples (190-191) provided by Müller, and then (192-193) provided by Naumkin & Porkhomovsky.

In Müller’s texts, the alleged use of *mon* as ‘what?’ is exceedingly rare, with (190) and (191) being the only examples found. Examples (190a-c) come from one single text on greeting formulae. These two examples cannot be explained away by appealing to dialectal differences with texts where *mon* appears in its regular meaning ‘who?’, because they come from the same speaker, a fisherman from East Soqotra (Müller 1905:vi-vii). Vladimir Agafonov (p.c.) has suggested that Müller might have misinterpreted (190a). According to him, the verb *baʕar* means ‘to say good morning’ and the suffix *-š* may stand for the 2SG.F.SBJ and not 2SG.F.OBJ, while the phrase *ʔl baʕer* is a fixed expression meaning ‘good morning!’. Consequently, (190a) may be translated as ‘Whom did you.F.SG say good morning?’. This interpretation may be perfectly acceptable when the sentence is considered in isolation, but I find it rather difficult to reconcile it with the overall context and with the existence of similar examples (190b) and particularly (190c). As to example (191), it has been confirmed by Vladimir Agafonov, who says to have encountered similar examples. Note, however, that *mon* in (191) may also be translated as ‘which one?’, i.e. ‘(Of the months that a pregnancy lasts) which is the month that you are now in?’. It is interesting to mention in this respect that in Shehri, another Modern South Arabian language, a cognate interrogative *mun* ‘who?’ is also found in *mun mən* ‘which of?’ (Simeone-Senelle 1997:390).

As far as Naumkin & Porkhomovsky’s examples (192) and (193) are concerned, it turns out to be possible to interpret *ínem* in both examples as

‘what?’ rather than ‘who?’ when the context and semantics are considered in more detail. Thus, example (192) comes from an archaic song and is uttered by a girl to her sheep/goat *Ṭawī:se*, when the two are followed by a spirit. That is, the use *nim* ‘what?’ here can be interpreted as an indication that the girl suspects that it is not a human being that follows them. As to (193), Naumkin & Porkhomovsky’s gloss of *ʾinim* ‘what?’ as Russian *kto* ‘who?’ here is simply because *čto* ‘what?’ is impossible in this context in Russian. This Soqotri example is rather an instance of ‘what?’-dominance in a KIND-question (as is not uncommon in Semitic, see Section III.2.3.4.1.3.2), where Russian would opt for ‘who?’-dominance. The normal answer to such a question would be to say how many boys and girls are left, rather than give the names of the children. In this respect, consider also (194) and its original German translation using *was* ‘what?’ and not *wer* ‘who?’.

Soqotri (Müller 1905:58)

- (194) *iném bīr-oh* *dī-é: bīoh?* [*ʿémor férihim*]
 what give.birth.to.PRF-3SG.F.SUBJ of-2SG.M mother
 ‘[And he asked him,] What has your mother given birth to? (‘Was hat deine Mutter geboren?’) [He said, ‘A girl.’]’

Summarizing, it appears that only Müller’s example (190), and perhaps also (191), may represent the real cases of lack of differentiation between ‘who?’ and ‘what?’, with *mon* ‘who?’ used as ‘what?’ instead of *iném* ‘what?’.

As example (191) may suggest, the semantic shift of *mon* in Soqotri has probably proceeded like what happened to *man* in Southern Mesopotamian Arabic dialects (see Section III.2.3.4.2.1). In particular, there may have existed an interrogative pronominal ‘which one? (person or thing)’. This interrogative then expanded its semantics to cover the meaning ‘who?’, thus entering in competition with the original interrogative ‘who?’ *mon* and causing some oscillation in the semantics of the latter. However, eventually, unlike in Southern Mesopotamian Arabic, it lost to *mon* and disappeared,¹¹⁴ leaving us with examples like (190) and (191). This disappeared interrogative ‘which one?’ was probably cognate to the Classical Arabic *ayy-* ‘which [N]?, what (kind of) [N]?’.

¹¹⁴ Remarkably, the sources on Modern South Arabian languages I consulted do not mention any attributive interrogative ‘which [N]?, what (kind of) [N]?’. Even in such relatively large collections of texts as Müller (1905) and Bittner (1917) I have not encountered any example with such an interrogative. It is possible of course that I have simply overlooked such an attributive interrogative. Most texts are not glossed and translations are sometimes quite free. Otherwise, this interrogative may exist but is particularly dispreferred and therefore absent from the texts. These languages then prefer a different strategy to ask for the same kind of information.

Its traces may be found in the interrogative ‘what?’ *īnem* < **ayna-m(a)* (see above in the present Section). Furthermore, the hypothesis proposed is indirectly supported by the following two facts from two other Modern South Arabian languages, Mehri and Shehri (or Jibbali). According to Müller (1905:373), in Mehri the meaning ‘who?’ is expressed either by *mo:n* or *éy*, the latter form being obviously cognate to the widespread Semitic interrogative ‘which [N]?, what (kind of) [N]?’. For Shehri, Simeone-Senelle (1997:390) gives *mun* ‘who?’ but also *mun mən* ‘which of?’.

2.3.4.2.3 The “Canaano-Akkadian” mixed language

According to Izre’el (1998:1), the “Canaano-Akkadian” mixed language is an extinct mixed East-West Semitic idiom which emerged in the second half of the second millennium BC in the land of Canaan (modern Israel, Lebanon, Palestine, and the western parts of Jordan and Syria) for the “upper diglossic medium” as a result of detachment of this region from the Akkadian “mainland” in the east after Canaan’s conquest by Egypt. Akkadian, an East Semitic language, was a lingua franca in Near East of the times, while Canaanite, a West Semitic language, was the indigenous language of Canaan. The data on this mixed language primarily come from the letters sent from Canaan to Egypt, which date approximately to the 14th century BC. In these letters, “Akkadian lexicon almost entirely predominated, but the grammar [...] was almost completely predominated by Canaanite [...] the mother-tongue of the scribes” Izre’el (1998:1). All in all, a little fewer than 300 texts and fragments of Canaano-Akkadian have been found.

The Canaano-Akkadian texts are “characterized by inherent variation” which “manifests itself in the form and meaning of both the Akkadian and the Canaanite components of the language, as well as in the employment of either component within the linguistic structure” (Izre’el 1998:3). In the case of interrogative pronominals, this variation makes interpreting the available data particularly difficult. As can be seen in (195), two forms, *mi:nu* and *mannu*, are attested each with two opposite meanings.

(195) *Canaano-Akkadian interrogative pronominals (mixed East-West Semitic idiom; ca. 1350 BC; Izre’el 1998:22-23)*

‘what?’ *mi:nu*, *mannu*

‘who?’ *mi:nu* (in Byblos), *mannu*, *miya*, *miyati*

The situation with the form *mi:nu* is somewhat clearer because it is attested with the meaning ‘who?’ only in “some letters” from Byblos, while *mannu* in letters from Byblos is used as ‘what?’ (Izre’el 1998:23). However, as far as the form *mannu* is considered, it is not clear from Izre’el’s description whether it was used

interchangeably as ‘what?’ and ‘who?’ by particular scribes or whether it was used as ‘what?’ by some scribes and as ‘who?’ by others, which, when all the texts are considered together, could give an impression of a ‘what?, who?’ variation.

In Akkadian proper, *mi:nu* meant ‘what?’, while *mannu* meant ‘who?’; compare Old Akkadian *min* ‘what?’ and *man* ‘who?’ and later in Old Babylonian *mi:nu(m)* (NOM)/ *mi:na(m)* ~ *mina:* (ACC) ‘what?’ and *mannu(m)* (NOM)/ *manna(m)* (ACC) ‘who?’ (East Semitic; Iraq; ca. 2350-2200 BC and ca. 2000-1500 BC, respectively; Lipiński 1997:328-329; Izre’el & Cohen 2004:48, 71, 100). The ‘who?’ interrogatives *miya* and (rare) *miyati* are purely Canaanite. The interrogative *mannu* used as ‘what?’ “reflects the cognate Canaanite form” (Izre’el 1998:23); compare Old Canaanite *manna* ‘what?’ (Lipiński 1997:328-329) and a rare early Biblical Hebrew (a later Canaanite idiom) form *mən* ‘what?’ (Gesenius & Kautzsch 1962:119; Macuch 1969:254).

If *mannu* was indeed used both as ‘what?’ and ‘who?’ by some scribes, this is surely a result of an imperfect fusion of Akkadian superstrate with Canaanite substrate (cf. Izre’el 1998:22), where the same or two very similar forms had opposite meanings. In all probability, the Byblos Canaano-Akkadian opposition *mi:nu* ‘who?’ vs. *mannu* ‘what?’ appeared due to a similar form-meaning conflict between the super- and substrate, but resolved in a slightly different way: the form *mannu* preserved its Canaanite meaning ‘what?’, while Akkadian *mi:nu* ‘what?’ was reanalyzed as ‘who?’ by analogy with the purely Canaanite ‘who?’ interrogatives, *miya* and *miyati*. Indirectly, the possibility of such reanalysis would also suggest morphologically complex origins of these interrogatives, with the first element *m(i)-* being the original interrogative element and the rest a “reinforcing” element, probably a deictic/copula (cf. Section III.2.3.4.1).

2.3.5 Cushitic languages

The Cushitic branch of Afro-Asiatic comprises four major sub-branches, Northern, Eastern, Central, and Southern. The former “West Cushitic”, now called Omotic, is usually considered to be a family distinct from Cushitic, sometimes even as not a part of Afro-Asiatic at all (cf. Bender 2000:1-2 for an overview). Some doubts have also been raised on the status of the Northern branch (cf. Lieberman 1990:566), represented by a single language Bedawi, which may turn out to be distinct from Cushitic. I will begin this Section by a general overview of the Cushitic interrogative pronominals in Section III.2.3.5.1. In Section III.2.3.5.2, I will examine the interrogative pronominals of the East Cushitic language Saho, which appears to be the only Cushitic language allowing for an indiscriminate use of one and the same interrogative pronominal as ‘who?’ or ‘what?’.

2.3.5.1 Cushitic interrogative pronominals: an overview

The interrogative pronominals of various Cushitic languages are summarized in Table 2. As far as happened to be possible, I also give here the forms of the attributive interrogative ‘which [N]?, what (kind of) [N]?’ which often seems to be capable of functioning pronominally as ‘which one?’ as well.

As can be observed, the ‘what?’ interrogative pronominals are mostly based on *ma* (or the like), or rarely *mi*, and a root meaning ‘thing’. For the meaning ‘who?’ Dolgopolsky (1991:12) reconstructs a Proto-Cushitic¹¹⁵ form **ʔayyu*, which is clearly related to the widespread Semitic interrogative *ʔayy-* (or the like) ‘which [N]?, what (kind of) [N]?’ (cf. Section III.2.3.4.1.1), Omotic **ʔay* ‘who?’ (primarily South Omotic) and **ʔay-(b-)* ‘what?’ (primarily North Omotic; Bender 2000:208-209).¹¹⁶ The Semitic meaning of this interrogative seems to be etymologically the oldest one, because it allows us to account much more easily for a further development both to ‘who?’ and to ‘what?’. Semantically, it also squares much better with the fact that this interrogative is clearly related to such forms as Proto-Semitic **ʔay* ‘where?’ (Lipiński 1997:328) and Dolgopolsky’s (1991:12) Proto-Cushitic **ay(y)u-da* ‘where?’.

In some languages, the shift from ‘which one?’ to ‘who?’ seems to have happened more than once. For instance, this seems to have happened in Northern Somali. While in Somali of Mogadishu ‘who?’ is *ya:* or *ayo* (Mauro Tosco, p.c.), which is a regular East Cushitic form, in Northern Somali it is *kúmà* in the masculine and *túmà* in the feminine (Dolgopolsky 1991:117; Mauro Tosco, p.c.). The interrogative ‘who?’ in Northern Somali distinguishes gender, as is typical for ‘which (one)?’ but not for ‘who?’ interrogatives. Moreover, it is formally quite similar to the ‘which (one)?’ interrogatives in other East Cushitic languages, e.g. in Oromo of Borana.

¹¹⁵ In fact, “Proto-Cushomotic”, i.e. both Proto-Cushitic and Proto-Omotic, since Dolgopolsky (1991) considers Omotic as West Cushitic, a part of a larger Cushitic family. The same remark is applicable elsewhere to all forms labelled as *Proto-Cushitic* with reference to Dolgopolsky (1991).

¹¹⁶ Somewhat similarly to the North Omotic form, I would have rather thought of the Proto-Cushitic interrogative as something like **ʔayya-(b(V)-)*, since quite a few Cushitic languages lack a labial in this interrogative. The final labial(ized) element, *-b(V)*, *-w(V)*, *-kʷ(V)* or the like, is likely to be a gender/case marker, like for instance the Bedawi masculine accusative marker *-b-* (North Cushitic; Sudan; Hudson 1976b:107).

Table 2. Some of the Cushitic interrogative pronominals (based on Appleyard 1987:264; Black 1976:231; Bliese 1976:151; Dolgopolsky 1991:117; Gragg 1976:179; Hudson 1976a:259; Hudson 1976b:128; Mous 1993:119; Sasse 1976:207; personal communication from Joachim Crass, Tolemariam Fufa, Maarten Mous, Graziano Savà, Mauro Tosco, Martine Vanhove)

Language	‘who?’	‘what?’	‘which (one)?’
North Cushitic			
Bedawi (Beja)	^ʔ <i>aw-</i> (NOM) ^ʔ <i>a:b-</i> (ACC)	<i>na:-n(a)</i> (<*‘what-thing?’)	<i>na:-</i> (? <i>na:-t</i> ‘thing’)
East Cushitic			
Burji	<i>ay(y)e</i>	<i>miya</i>	
Gedeo (Darasa)	<i>ayye</i>	<i>ma:</i>	
Hadiyya	<i>ayye</i>	<i>maha</i>	
Kambata	<i>aye</i>	<i>ma</i>	
Libido	^ʔ <i>aye</i>	<i>maha</i> <i>mahicco</i> (SG.ACC)	<i>hinki</i>
Sidamo	<i>ay(e)</i>	<i>ma(:)/ma(y)e/ma(y)</i>	<i>hikko</i> (M.SG) <i>hitte(na)</i> (F.SG)
Afar	^ʔ <i>yya</i>	<i>ma</i>	
Oromo of Wellegga	<i>eeɽɽu(-tu)</i> (NOM)	<i>maaltu</i> (NOM)	<i>kamtu</i> (NOM)
Oromo of Borana	<i>é:n(n)u</i>	<i>má:ní</i>	<i>kámí</i> (M.SG) <i>támí</i> (F.SG)
Daasanach (Dasenech)	<i>mááya</i>	<i>méétte</i>	
Gawwada	<i>yáfa</i>	<i>móʔo</i>	
Tsamai	^ʔ <i>aha</i>	<i>mo</i>	<i>kunda</i>
Somali of Mogadishu	<i>ya:/ayo</i>	<i>maaxa</i>	<i>ce:</i> (M.SG)
Northern Somali	<i>kúmà</i> (M) <i>túmà</i> (F)	<i>maahay</i> (<*‘what-thing?’)	<i>kɛ:</i> (M.SG) <i>tɛ:</i> (F.SG)
Central Cushitic			
Awngi	^ʔ <i>ayí</i>	<i>wátay</i>	
Bilin	<i>aw-</i>	<i>war-a:</i> (<*‘what-thing?’)	
Xamtanga	<i>aw</i> (NOM)	<i>wár/wǎǎǎǎ</i>	
South Cushitic			
Iraqw	<i>he:-m-á</i> (<*‘man-what-be?’)	<i>milá/m-</i>	<i>ga:lá</i> (<i>ga:</i> ‘thing’)

2.3.5.2 Saho: an East Cushitic language with a ‘who?, what?’ interrogative

The East Cushitic language Saho, as described by Reinisch (1878, 1889, 1890), appears to be the only Cushitic language allowing for an indiscriminate use of one and the same interrogative pronominal as ‘who?’ or ‘what?’. The use of a single ‘who?, what?’ interrogative is typical for all Saho dialects, except that of Asa-Awurta (Asaorta) and Tarua (Taru‘wa; Reinisch 1890:253). Saho interrogative pronominals are summarized in Table 3. For convenience sake, I will refer to the Saho described by Reinisch as *Non-Asa-Awurta Saho* further on.

Table 3. Saho interrogative pronominals

	Asa-Awurta Saho & Tarua Saho (Conti Rossini 1913; Reinisch 1890)	Other Saho varieties (Reinisch 1890)
‘who?’	Reinisch (1890:253): <i>ma:/mi:</i> Conti Rossini (1913:6-8): <i>mi</i> <i>mi-yattú ~ mi[?]attú ~ mi[?]attí ~</i> <i>me-yattò</i>	Reinisch (1890:2, 56): <i>ay</i> <i>a:-tí:ya:</i> (M; <- <i>tí:ya:</i> ‘one.M’) <i>á:-tya:</i> (F; <- <i>tyá:</i> ‘one.F’) <i>a:-má:ra:</i> (PL; <- <i>má:ra:</i> ‘those who live’)
‘what?’	Conti Rossini (1913:6-7): <i>ay ~ a</i> <i>ay-ím ~ a-[?]ím ~ ay-mì</i>	Reinisch (1890:2, 56): <i>ay ~ a:</i>
‘which [N]?, what (kind of) [N]?’	Conti Rossini (1913:6-7): <i>a</i>	Reinisch (1890:2): (<i>a: ~ ay</i>)
‘which one?’ (thing)	Conti Rossini (1913:7): <i>a-tiya:</i> (M; <- <i>tíya:</i> ‘one.M’) <i>a-ti[?]á:</i> (F; <- <i>tiyá:</i> ‘one.F’)	

Conti Rossini (1913) does not speak about any functional difference between forms like *mi* and *mi-yattú*, *mi-yattí*, *me-yattò*, on the one hand, and *ay* and *ay-ím*, on the other. Nor does he say anything on the origin of the second parts in the more complex forms.¹¹⁷ Examples (196-204) illustrate the uses of the interrogatives in Asa-Awurta Saho as found in Conti Rossini (1913).

¹¹⁷ The first one, *-yattú*, may be related to the Libido singular accusative marker *-yicco* in *ma-yicco* ‘what?’. It may also be composed of *ya* ‘be.3SG.IPRF’ (Conti Rossini 1913:17) and a nominalizing/singulative suffix *-(t)ta*, *-(t)to*, *-(t)tu*, *-(t)tá:*, *-(t)tó:*, as in *bolol-tò* ‘burnt, the burnt one’ from *bolol* ‘to burn’, *baryatto* ‘one slave’ from *barya:* ‘slave’, *olal-to* ‘one/a euphorbia plant’ from *olal* ‘euphorbia plant(s)’ (Conti Rossini 1913:23, 25). The element *-ím ~ -mì* may be the same as the formative *-m ~ -ím* used to construct independent possessive

Asa-Awurta Saho

- (196) *mi(-²attú) rabé?*
 IPW die.PR.F.3SG.SBJ
 ‘Who died?’ (Conti Rossini 1913:6-7)
- (197) *mi faràs kinní?*
 IPW horse exist.IPR.F.3SG.SBJ
 ‘Whose horse is it?’ (Conti Rossini 1913:6)
- (198) *ta: numa: meyattò kinní?*
 this woman IPW exist.IPR.F.3SG.SBJ
 ‘Who is this woman?’ (Conti Rossini 1913:7)
- (199) *ay kinní ta:?*
 IPW exist.IPR.F.3SG.SBJ this
 ‘What is this?’ (Conti Rossini 1913:6)
- (200) *a’ím tobbé?*
 IPW hear.IPR.F.2SG.SBJ
 ‘What have you heard?’ (Conti Rossini 1913:6)
- (201) *a-tí’á: kinní ka-‘aré:?*
 IPW-one.F exist.IPR.F.3SG.SBJ 3SG.M.POSS-house
 ‘Which one is his house?’ (Conti Rossini 1913:6)
- (202) *tamá: a numá:?*
 this IPW woman
 ‘Who is this woman? (lit.: ‘What woman is this?’)’ (‘costei che donna è? = chi è questa donna?’), Conti Rossini 1913:6)
- (203) a. *aymì ka-migá: kinní?*
 what 3SG.M.POSS-name exist.IPR.F.3SG.SBJ
 ‘What is his name?’ (Conti Rossini 1913:6-7)
- b. *ta heyaw-tí miyattò ka-migá:?*
 this man-GEN who 3SG.M.POSS-name
 ‘What is the name of this man?’ (Conti Rossini 1913:32)
- (204) *ta dik ay-nú: ka-migá: kinní?*
 this village what-? 3SG.M.POSS-name exist.IPR.F.3SG.SBJ
 ‘What is the name of this village?’ (Conti Rossini 1913:6)

Example (202) appears to represent an attributive use of *a*, as ‘which [N]?, what (kind of) [N]?’.¹¹⁸ For the collocation *a-tiya:* (M)/*a-tí’á:* (F) ‘which one?’ Contini

pronominals, as in *yí-m* ‘mine’ from *yí* ‘my’ (cf. Conti Rossini 1913:5-6).

¹¹⁸ Conti Rossini (1913:7) also gives the form *ay-ddá:*, which he glosses as ‘which? (quale?)’, although without providing an example, but elsewhere (1913:31) as ‘how much/many? (quanto?)’.

Rossi gives only an example with a question about a thing (201) while Reinisch (1890) gives examples only with questions about persons (see below). Note that questions about proper names may involve ‘who?’-dominance when the question is about a personal proper name, as in (203b), but apparently not when the question is about a place-name, as in (204).

Non-Asa-Awurta Saho, as described by Reinisch, has only one interrogative pronominal *ay*, which may also surface fused as *a:*, emphasized as *ejí:*, and in the genitive as *eyí:*, *eyí:* or *ejí:* (Reinisch 1890:2, 56).¹¹⁹ Its pronominal use as ‘who?’ is illustrated in (205) and (206), and as ‘what?’ in (207-209).

Non-Asa-Awurta Saho

- (205) *ta: heyôto: ay du:dúsa?*
 this man IPW fool.IPRF.3SG.SBJ
 ‘Who is fooling this man?’ (Reinisch 1890:56)
- (206) *ejí: bá:la: kí:to:?*
 IPW.GEN son exist.IPRF.2SG.SBJ
 ‘Whose son are you?’ (Reinisch 1890:56)
- (207) *atú: ta-m ay?*
 2SG say.IPRF.2SG.SBJ-REL IPW
 ‘What do you say? (lit.: ‘What (is it) that you say?’) (Reinisch 1890:56)
- (208) *atú: bé:tta-m a: lí:to:?*
 2SG eat.IPRF.2SG.SBJ-REL IPW at-2SG.GEN
 ‘What do you have to eat? (lit.: ‘What (is it) that at you (is that) you eat?’)
 (Reinisch 1890:2)
- (209) *ay fał-da, yi ba:lá:û?*
 what want.IPRF.2SG.SBJ 1SG.POSS daughter.VOC
 ‘What do you want, my daughter?’ (Reinisch 1889:161)

Depending on the syntactic analysis, examples (210) and (211) may be interpreted either as an instance of its attributive use, as ‘which [N]?, what (kind of) [N]?, or as examples of its pronominal use as ‘who?’ and ‘what?’, respectively.

Non-Asa-Awurta Saho

- (210) *a: nu:má: tametá-tiya:?*
 IPW woman come.PRF.3SG.SBJ-one.F
 ‘What/which woman has come?’ (Reinisch 1890:2) or ‘Who is the woman who came?’

¹¹⁹ The use of a special genitive form seems to be optional.

- (211) *a: ʔa:‘emó: sí:na: la?*
 IPW taste 2PL at.3SG.GEN
 ‘How do you like the taste? (lit.: ‘What taste (do) you have at it?’ or
 ‘What (is) the taste you have at it?’)’ (Reinisch 1890:359)

The form *a:* is also used in collocation with the word ‘one’ as *a:-tí:ya:* (M) and *á:-tya:* (F), in the singular, and in the plural with a deverbal noun *má:ra:* ‘people, inhabitants (those who live, who dwell)’ (from *ma:r* ‘to live, to dwell’). For these collocations Reinisch (1890) gives only examples where they mean ‘who?’, (212-213). This contrasts with Conti Rossini’s (1913) example in (201), where such a collocation is used as ‘which one? (thing)’.

Non-Asa-Awurta Saho

- (212) *atú: a:-tí:ya:?*
 2SG IPW-one.M
 ‘Who are you?’ (Reinisch 1890:2)
- (213) *ta: nu:má: á:-tya:?*
 this woman IPW-one.F
 ‘Who is this woman?’ (Reinisch 1890:2)

As readily follows from the discussion in Section III.2.3.5.2, the original meaning of the Saho interrogative *ay* must have been the pronominal ‘which one?’ and/or attributive ‘which [N]?, what (kind of) [N]?’. Subsequently, it was extended to non-selective contexts as ‘who?’ and ‘what?’.

Comparison with the interrogatives of other East Cushitic languages, and especially with those of Afar, which is the East Cushitic language most closely related to Saho (see Table 2), suggests that the ‘who?’ interrogative *mi:/ma:* of Asa-Awurta and Tarua Saho is either an extremely archaic feature or rather represents a recent innovation. In the latter case, it is in all probability a loan from a neighbouring Semitic language, such as Tigrinya *mān* ‘who?’, or a Modern Arabic variety, such as the Yemenite Arabic *man/min* ‘who?’ (see Table 1 in Section III.2.3.4.1.1), with the loss of the final *n*.

3 Eurasia

In this section, I will discuss the languages presumably allowing a lack of differentiation between ‘who?’ and ‘what?’ that are spoken in Eurasia (which excludes Sino-Tibetan, Tai-Kadai, Austro-Asiatic,¹ and Andamanese languages of Southeast Asia). The section is organized as follows. In Section III.3.1, I will examine Indo-European languages, in Section III.3.2 North Caucasian languages, in Section III.3.3 Dravidian languages, in Section III.3.4 Kusunda, and in Section III.3.5 Tungusic languages.

3.1 Indo-European

3.1.1 Introduction

Indo-European is one of the largest linguistic groups in the world, both in the number of languages and the number of their speakers. For the reasons explained in Section III.1, it happened to be impossible to examine Indo-European languages in the same degree of detail as many of the other Old World languages. The only Indo-European languages commonly mentioned in the typological literature to allow a lack of differentiation between ‘who?’ and ‘what?’ are Latvian and Lithuanian (cf., e.g., Hjelmslev 1956; Maytinskaya 1969; Ultan 1978; Lindström 1995), the only two East Baltic languages (Section III.3.1.2). However, a certain lack of differentiation between ‘who?’ and ‘what?’ appears to be possible in quite a few other Indo-European languages as well.² In particular, in what follows, I will discuss data of French (Section III.3.1.3), Old English (Section III.3.1.4), Ancient Greek (Section III.3.1.5) and Latin (Section III.3.1.6). It should be pointed out that more Indo-European languages may eventually prove to allow for similar kinds of lack of differentiation between ‘who?’ and ‘what?’.

The presumed lack of differentiation between ‘who?’ and ‘what?’ in the Indo-European languages discussed is most common with predicatively used interrogatives. Another important factor that appears to have played a particularly favourable role in the development of a lack of differentiation between ‘who?’ and ‘what?’ is that in older Indo-European languages the interrogative pronominals basically distinguished gender-number, while the human vs. non-human distinction was subsequently determined by the semantics of genders and

¹ Note that although strictly speaking, Munda Austro-Asiatic languages belong to Eurasia, I will discuss them together with the other Austro-Asiatic languages in Section III.4.1.

² Note that I do not include here the lack of differentiation determined by case, which is much more common.

agreement patterns.

More specifically, the Indo-European languages discussed can be subdivided as follows with respect to the factors that may have contributed to the development of a lack of differentiation between ‘who?’ and ‘what?’: (i) the use of an interrogative pronominal differentiating gender-number rather than humanness/non-humanness and the peculiarities of the evolution of the gender-number system (East Baltic languages and Ancient Greek), (ii) the predicative use of an interrogative pronominal, differentiating gender-number rather than humanness/non-humanness, and peculiarities of the gender system of the language (Latin and Ancient Greek), (iii) predicative use of the otherwise attributive interrogative in itself indifferent to the opposition human vs. non-human (Modern French), (iv) the peculiarities of the way the older system of the interrogative and relative pronominals has been structurally reorganized and reduced in the daughter language (Middle and Classical French, some Modern French dialects and French-based creoles). I have not included here the presumed lack of differentiation between ‘who?’ and ‘what?’ with the Old English interrogative pronominal *hwæt*, since the situation is somewhat less clear there.

3.1.2 *East Baltic languages*

The Baltic family consists of two branches, East and West Baltic. The western branch includes several long-extinct languages, with Prussian being the best-known representative. The eastern branch comprises the only two living Baltic languages, Latvian and Lithuanian. Both Latvian and Lithuanian allow for lack of differentiation between ‘who?’ and ‘what?’. The two languages are also quite similar in both the forms of their interrogative pronominals and the ways these are used. Thus, Lithuanian has *kàs* ‘who?, what?’ and Latvian *kas* ‘who?, what?’, neither pronoun has a plural form. The declensions of the two pronominals are summarized in Table 1.

Table 1. Case paradigms of Lithuanian kàs ‘who?, what?’ (Ambrazas 1997:209) and Latvian kas ‘who?, what?’ (Holst 2001:131)

	Lithuanian	Latvian
NOM	<i>kàs</i>	<i>kas</i>
GEN	<i>kõ</i>	<i>kā</i>
DAT	<i>kám</i>	<i>kam</i>
ACC	<i>kā</i>	<i>ko</i>
INS	<i>kuõ</i>	
LOC	<i>kamè</i>	

The examples in (1) illustrate the use of the ‘who?, what?’ interrogative for Latvian.

Latvian

- (1) a. *kas tas ir?*
IPW.NOM DEM.M.SG.NOM is
‘Who/what is this?’ (Nau 1999:146)
- b. *kā te trūkst?*
IPW.GEN here miss.PRS.3
‘Who/what is missing here?’ (Nau 1999:146)
- c. *kam tas vajadzīgs?*
IPW.DAT DEM.M.SG.NOM necessary
‘Who needs it?’ or ‘Why is this necessary?’ (Nau 1999:147)
- d. *ko tu redzi?*
IPW.ACC 2SG.NOM see.PRS.2SG
‘Who/what do you see?’ (Nau 1999:146)

Besides the forms mentioned in Table 1, Lithuanian also has a special possessive interrogative *kienõ* ‘whose?’, as in (2). Such a possessive form is lacking in Latvian.

Lithuanian (Ambrazas 1997:193)

- (2) *kienõ tà knygà?*
whose DEM.F.SG.NOM book.F.SG.NOM
‘Whose book is it?’

According to Dambriūnas *et al.* (1972:285), the same form *kienõ* is also used to express the human subject of a passive construction. Dambriūnas *et al.* (1972) do not pronounce themselves explicitly on which form needs to be used when the subject of the passive is an interrogative pronominal with a non-human meaning. However, example (3) seems to suggest that the nominative *kàs* should be used.

Lithuanian (Dambriūnas *et al.* 1972:272)

- (3) *kàs yrà dāroma iš giñtaro?*
IPW.NOM is make.PRS.PASS.PTCP.N of amber.M.SG.GEN
‘What is made of amber?’

Besides the possessive form *kienõ*, Lithuanian has three other special possessive forms, viz. *màno* 1SG.POSS.SG vs. *manęs* 1SG.GEN, *tàvo* 2SG.POSS vs. *tavęs* 2SG.GEN, and the reflexive *sàvo* REFL.SG.POSS vs. *savęs* REFL.SG.GEN, which also must be used to “express the semantic subject in a passive construction” (Dambriūnas *et al.* 1972:192-193, 276-277). The existence of a special possessive form *kienõ* ‘whose?’ vs. the genitive *kõ* does not affect in any way the

fact that the lexeme represented in the nominative by the form *kàs* in Lithuanian is as such indifferent to the opposition human vs. non-human. To begin with, the possessive forms, such as *kienõ*, *manẽs*, etc., may be considered as distinct lexemes, rather than special genitive case forms of respectively *kàs*, *màno*, etc.³ Furthermore, even if *kienõ* were treated as a second case form of the nominative *kàs*, its human meaning could be accounted for as conditioned by the semantics of the possessive.⁴

In both Latvian and Lithuanian, various strategies exist for conveying the meanings ‘who?’ and ‘what?’ in a less ambiguous way. For instance, in colloquial Latvian the masculine singular form of the selective interrogative *kurs* ‘which (one)?’ is often used non-selectively as the human interrogative ‘who?’ (Holst 2001:131; Nau 1999:145-147; Nau 1999:145-147 also discusses other strategies for the disambiguation between the meanings ‘who?’ and ‘what?’ in Latvian).

By way of conclusion, a few words can be said on the origins of the lack of differentiation between ‘who?’ and ‘what?’ in Latvian and Lithuanian. As discussed by Holst (2001:131), the reason for this lack of differentiation lies in the peculiarities of the evolution of the gender systems of the East Baltic languages. Thus, in older Indo-European languages, the meaning ‘who?’ is basically expressed with the masculine form of the interrogative pronominal,⁵ while the meaning ‘what?’ with the neuter form of the same interrogative pronominal. However, in East Baltic languages the system of nominal gender has been reorganized in such a way that the masculine fell together with the neuter.⁶ Lithuanian *kàs* and Latvian *kas* ‘who?, what?’ are originally the masculine forms of the interrogative pronominal.⁷ Compare, in this respect, the Prussian (West

³ This is, for instance, how similar forms in Slavic languages are normally treated, e.g. Russian *čej* ‘whose (M.SG.NOM)?’ (vs. *kogo* ‘who.GEN?’ or *čego* ‘what.GEN?’), *moj* ‘my (M.SG.NOM)’ (vs. *menja* 1SG.GEN), etc. The only major difference between the Russian and the Lithuanian possessives seems to be the fact that in Russian, possessives show agreement in gender, number and case with the possessed noun, e.g. *č’ej* ‘whose(F.SG.GEN)?’.

⁴ It may be interesting to point out in this respect that the Russian possessive interrogative *čej* ‘whose?’ mentioned in the preceding footnote is also restricted to questions about the same range of entities as the regular interrogative *kto* ‘who?’, i.e. about humans and (larger) animals.

⁵ There is sometimes a distinct feminine form of the interrogative pronominal as well, but it is always the masculine form that appears to be used by default, i.e. when the sex or number of person(s) is not unknown or irrelevant.

⁶ There are some relics of neuter gender in adjectives and deictics, which are used for purposes of enforced agreement and similar contexts (cf. Ambrazas 1997:134-137).

⁷ It may be interesting to note that at least in Latvian, this originally masculine form, which used to be the default form in questions about humans and thus the closest equivalent of ‘who?’, but

Baltic) interrogative pronominals: M.SG *kas* (also the default ‘who?’), F.SG *kā*, N.SG *ka* ‘what?’, and PL common to all genders *kai* (Endzelin 1923:398; Mažiulis, ms. [2004]:75; Schmalstieg 1974).

3.1.3 French

In the history of French, there appear to be two instances of a certain lack of differentiation between ‘who?’ and ‘what?’. The first case, discussed in Section III.3.1.3.1, is represented by the Modern French interrogative *quel*. The second case, discussed in Section III.3.1.3.2, involves the interrogative *qui*, which in Modern French means only ‘who?’, but used to also mean ‘what?’ in Middle and Classical French and later in some French dialects and French-based creoles.

3.1.3.1 Modern French *quel*?

Modern French has three simple interrogative pronominal forms, one human *qui* ‘who?’ and two non-human *que* and *quoi* ‘what?’. The first of the two non-human forms, *que*, is mostly restricted to the preverbal direct object function, while *quoi* is most common postverbally, as both subject and object, and after prepositions. Besides the three simple forms, French makes extensive use of four periphrastic constructions based on *qui* and *que*, viz. *qui est-ce qui* ‘who? (NOM)’ (lit.: who is-it REL.NOM), *qui est-ce que* ‘who? (ACC)’ (lit.: who is-it REL.ACC) and *qu’est-ce qui* ‘what? (NOM)’ (lit.: what is-it REL.NOM), *qu’est-ce que* ‘what? (ACC)’ (lit.: what is-it REL.ACC). Furthermore, French has an interrogative *quel* ‘which [N]?, what (kind of) [N]?’,⁸ usually used attributively, as in *quel arbre/homme?* ‘which tree/man?, what (kind of) tree/man?’, and its derivative *lequel* ‘which one?’. However, *quel* is also used predicatively and then it can inquire about both a human and a non-human subject, as in (4) and (5) respectively (Riegel *et al.* 2001:395).⁹ That is, *quel* means ‘what?’ in (4) and ‘who?’ in (5).¹⁰

now functions as both ‘who?’ and ‘what?’, may one day end up as a dedicated non-human interrogative ‘what?’, provided the use of the masculine singular form of the selective *kurš* ‘which one?’ in the meaning ‘who?’ described above becomes the norm (cf. Holst 2001:131).

⁸ This is the masculine singular form. Other forms are F.SG *quelle*, M.PL *quels* and F.PL *quelles*. Note, however, that all the four forms are normally pronounced the same, viz. /kɛl/.

⁹ I am grateful to Peter Lauwers (p.c.) for suggesting this reference.

¹⁰ The predicative use of *quel* to inquire about the quality or the kind of subject has been common till the 19th century but it is archaic in contemporary French (Riegel *et al.* 2001:395).

French

- (4) a. *quel est cet arbre?*
 IPW.M.SG is DEM.M.SG tree[M.SG]
 ‘What tree is this/that? (lit.: ‘What is this/that tree?’)’
- b. *quel est le titre de ce texte?*
 IPW.M.SG is the.M.SG title[M.SG] of DEM.M.SG text[M.SG]
 ‘What is the title of this text?’
- (5) a. *quel est cet homme?*
 IPW.M.SG is DEM.M.SG man[M.SG]
 ‘Who is this/that man?’ (Riegel *et al.* 2001:395)
- b. [A: *Vous me conterez tout cela. Je m’attends bien à du nouveau; mais en vérité je n’en veux pas encore. Comme ce lavoir est petit! autrefois il me paraissait immense; j’avais emporté dans ma tête un océan et des forêts, et je retrouve une goutte d’eau et des brins d’herbe.*
Quelle est donc cette jeune fille
 IPW.F.SG is PTCL DEM.F.SG young.F.SG girl[F.SG]
 [qui chante à sa croisée derrière ces arbres?]
 [B: *C’est Rosette, la soeur de lait de votre cousine Camille.*]
 ‘[A: You should tell me all this. I’m really expecting something new, but to tell the truth, I don’t want it yet. How small this wash-house is! In the old days, it looked so huge. In my head, I had brought an ocean and woods, and I find a drop of water and blades of grass.] Who is in fact that girl [who is singing in her window behind these trees? B: It’s Rosette, a foster-sister of your cousin Camille]’ (Alfred de Musset *On ne badine pas avec l’amour*, reference via Rey-Debove *et al.* 1996:1837)

The use of predicative *quel* as ‘what?’ in sentences like (4a) and (4b) is the norm and other (non-human) interrogative pronominals are hardly possible here. Examples in (5) are somewhat different in that one can just as well use the ordinary human interrogative pronominal *qui* ‘who?’ instead, as in (6).

French

- (6) *qui est cet homme?*
 who is DEM man
 ‘Who is this/that man?’ (Riegel *et al.* 2001:395)

Furthermore, *qui* ‘who?’ becomes the only option in this construction when the subject is a personal pronoun (Riegel *et al.* 2001:395), as in (7).

French

- (7) *qui êtes-vous, vous qui m'écoutez?*
 who are-2PL 2PL REL.NOM me-listen.2PL
 'Who are you who are listening to me?' (Riegel *et al.* 2001:395)

Similarly, predicative *quel* as 'what?' is not possible if the subject is a deictic, such as *ce* 'it, this, that'. Instead, *qu'est-ce que* 'what?' (8a) or more informally, *quoi* 'what?' (8b) need to be used.

French

- (8) a. *qu'est-ce que c'est?*
 what-is-DEM REL.ACC DEM-is
 'What is this/that?'
 b. *c'est quoi?*
 DEM-is what
 'What is this/that?'

Finally, it is worth mentioning here that according to Riegel *et al.* (2001:395), till the 19th century it was possible to use a predicative *quel* with respect to persons in the case of a non-prototypical combination of values of the kind [person + classification (+ common noun)] (instead of *que* or *quoi* 'what?', cf. Section I.4.2.3.1). Given that *quel* goes back to Latin *quālis* 'what kind of [N]?', this classificational use of predicative *quel* with respect to persons must have developed before its identificational use. In this respect, it may be also interesting to compare *quel* with the Ancient Greek interrogative *poios* 'what (kind of) [N]?' (sometimes also 'which [N]?'), which in Modern Greek has developed into a full-fledged interrogative 'who?' (cf. Section III.3.1.5).

3.1.3.2 Middle and Classical French, non-standard French varieties and French-based creoles: the interrogative *qui*?

Another instance of lack of differentiation between 'who?' and 'what?' attested in the history of French involves the interrogative pronominal *qui*. Although in Modern French *qui* means only 'who?', from around the 13th century until around the 17th-18th centuries, i.e. from the Middle French period until the end of the Classical French period, it was also used as 'what?' next to the dedicated 'what?' interrogatives *que* and *quoi* (cf. among others, Bloch & von Wartburg 1968:525; Darmesteter 1902:III, §§416-418; Grévisse 1986:§701; Nyrop 1935:V, §356).¹¹ After this period the use of *qui* as 'what?' can be found only in some

¹¹ In literary and poetic styles of Modern French, the non-human use of *qui* can still be sporadically attested even in the 20th century (Grévisse 1986:§701).

non-standard French varieties and many French based creoles.

According to some sources (e.g., Nyrop 1935:V, §356), the non-human use of *qui* had been generalized only in Middle French of the 15th century. Two examples in (9) provide an illustration of *qui* used in the meaning ‘what?’

Classical French

- (9) a. *Qui fait l’oiseau? c’est le plumage*
 IPW makes the-bird DEM-is the feathers
 ‘What makes a bird? It is the feathers’ (La Fontaine *La chauve-souris et les deux belettes* via Darmesteter 1902:III, §§416)
- b. *Après ce coup, Narcisse, à qui dois-je m’attendre?*
 after DEM blow PROP to IPW must-1SG 1SG.REFL-wait.for
 ‘After this blow, Narcissus, what must I expect?’ (Racine *Britannicus* via Darmesteter 1902:III, §§416)

The use of *qui* as ‘what?’ appears to have been most common in the subject function, as in (9a). Some sources, such as Nyrop (1935), even claim that *qui* as ‘what?’ was restricted to the subject function, but this is contradicted by example (9b), where *qui* is used in the meaning ‘what?’ as the complement of a preposition. The use of *qui* ‘what?’ as direct object appears to have been very rare, since the sources consulted remain silent on this possibility. (10) and (11) are the only two examples I found where *qui* is used as a direct object and the entity questioned about is a thing.

Classical French

- (10) *Je ne sçay qui je dois admirer davantage*
 1SG NEG know IPW 1SG must admire more
Ou de ce grand amour, ou de ce grand courage
 or of DEM big love or of DEM big courage
 ‘I do not know what I should admire more, this big love or this big courage?’ (Corneille *L’illusion comique* via Grévisse 1986:§701)¹²
- (11) *Qui puis-je faire?*
 IPW can-1SG do
 ‘What can I do?’ (via *Antidote* electronic dictionary,¹³ the original source is not indicated)

Furthermore, according to Darmesteter (1902:III, §§418), till the 17th century *qui* was frequently used predicatively as a selective interrogative ‘which one(s)?’ instead of *quel*, as in (12).

¹² Admittedly, this is only an indirect question and the context here is rather selective.

¹³ © Druide informatique Inc. 1993-2001, www.druide.com.

Classical French

- (12) *Entre tant d'animaux, qui sont ceux qu'on estime?*
 among so.much of-animals IPW are DEM.PL REL.ACC-one esteems
 'Among the many animals, which are those that we esteem?' (Boileau
Satire via Darmesteter 1902:III, §§418)

Another apparent restriction on the non-human use of the interrogative *qui* in Middle and Classical French was that it was not used in the periphrastic constructions mentioned in the beginning of Section III.3.1.3.1. That is, there appear to be no examples of the periphrastic constructions *qui est-ce qui* and *qui est-ce que* used in the meaning 'what?' (Rottet 2004:172).

All in all, it seems that for at least a small fraction of the Middle and Classical (Standard) French authors there was no restriction on the use of *qui* as 'what?' in the direct object function, which implies that in their speech a genuine lack of differentiation between 'who?' and 'what?' was possible. However, the restriction on the use of *qui* as 'what?' in the direct object function appears to have been the norm (*que* or *quoi* were used instead). In other words, the normal situation was a partial neutralization of differentiation between 'who?' and 'what?' determined by case. As such, this would not be uncommon among Indo-European languages (cf. Section III.3.1.6 on Latin or Section III.3.1.5 on Ancient Greek). What would be special, however, is that (i) in Middle and Classical French, a distinct non-human form of the interrogative was still available in every syntactic function, whereas for instance, this was not the case in Latin (e.g., in the dative the Latin interrogative pronominal had only one form, viz. *cui*, cf. Table 3 in Section III.3.1.6), (ii) the lack of differentiation between 'who?' and 'what?' in Middle and Classical French existed primarily in the nominative, whereas in most other Indo-European languages, such as Latin, if the two meanings are expressed with one form, this form would normally belong to one of the non-nominative cases (cf. Table 3 in Section III.3.1.6).

The partial neutralization of differentiation between 'who?' and 'what?' determined by case with *qui* in Middle and Classical French has developed into a full lack of differentiation in some non-standard French varieties and many French-based creoles. In the following description of the distribution of *qui* 'who?', 'what?' in non-standard French, I base myself on Rottet (2004), where more examples and references can be found. The interrogative *qui* is (or used to be) found as the only general interrogative pronominal 'who?', 'what?' in a number of western French dialects, including those from Normandy, eastern Brittany, Poitou, Anjou, parts of Saintonge and Berry.¹⁴ In these dialects, there are no syntactic restrictions on the use of *qui*. It can also function as direct object

¹⁴ Some of these dialects have preserved the dedicated 'what?' interrogative *quoi* in a fossilized form in *pourquoi* 'why?' (lit.: 'for what') (cf. Rottet 2004:173).

and in periphrastic constructions, as in (13) from Poitevin (i.e. the French dialect of Poitou).

Poitevin French

- (13) a. *Qui qu'est venu?*
 IPW REL-is come.PST.PTCP
 'Who came?' (Mineau 1982:255 via Rottet 2004:173)
- b. *Qui qu'tu manges?*
 IPW REL-2SG eat.PRS.2SG
 'What are you eating?' (Mineau 1982:255 via Rottet 2004:173)
- b. *Mais qui est-ce qu'i vous est donc arrivé, Père Sanfaçon?*
 but IPW is-DEM REL-3SG.M 2PL is PTCL happened PROP
 'But what happened to you, Père Sanfaçon?' (Favreau *et al.* 1983:143 via Rottet 2004:173)

The expansion of *qui* as the single general interrogative pronominal 'who?, what?' at the expense of the dedicated 'what?' interrogatives *que* and *quoi* in western French dialects must have been completed already by the 17th century. This is suggested by the fact that many colonial French varieties, which appeared in the 17th-18th centuries and which must have had an important input from the western French dialects just mentioned, as well as many French-based creoles, which in turn evolved on the basis of these colonial French varieties, also use a single interrogative pronominal *qui* 'who?, what?' (and periphrastic constructions on its basis). For instance, this is the situation found in Louisiana French (more specifically in the varieties spoken in the lower Lafourche Basin in the southeast and in Evangeline and Avoyelles Parishes in the northwest) (14) and its offshoot Louisiana Creole (15).¹⁵

Louisiana French

- (14) a. *Qui c'est qui fait ce train?*
 IPW DEM-is REL.NOM makes DEM noise
 'Who/what is making that noise?' (Guilbeau 1950:169 via Rottet 2004:175)
- b. *Qui vous-autres aurait fait si j'avais pas de licence?*
 IPW 2PL-others would.have done if 1SG-had NEG of license
 'What would you all have done if I didn't have a license?' (Rottet 2004:176)

¹⁵ These idioms have preserved the dedicated 'what?' interrogative *quoi* only in some frozen collocations, such as *pourquoi* 'why?' (lit.: 'for what') and *quoi faire* 'how?' (lit.: 'what do'), although the same interrogatives based on *qui*, viz. *pourqui* 'why?' and *qui faire* 'how?', are also attested (cf. Rottet 2004:178, 184).

Louisiana Creole¹⁶

- (15) a. *Qui ça qui senti si bon dans la chaudière la,*
 IPW DEM REL.NOM smells so good in the kettle DEM
Compair Lapin?
 PROP
 ‘What smells so good in the kettle, Brer Rabbit?’ (Fortier 1895:30 via Rottet 2004:183)
- b. *Ki-sa ki nouzot kandida?*
 IPW-DEM REL.NOM 1PL candidate
 ‘Who is our candidate?’ (Neumann-Holzschuh 1987:99 via Rottet 2004:183)
- c. *Qui ça t’olé?*
 IPW DEM 2SG-want
 ‘What do you want?’ (Neumann-Holzschuh 1987:99 via Rottet 2004:183)
- d. *Ki sa to wa?*
 IPW DEM 2SG see
 ‘Who do you see?’ (Lane 1935:12 via Rottet 2004:183)

In Louisiana Creole, the original periphrastic construction, as in (15), is sometimes further truncated to *sa ki* ‘who?, what?’ in the nominative (16a, b) and just *sa* ‘who?, what?’ in the accusative (16c), where the form *sa* is the French neuter demonstrative pronominal *ça*. The accusative form *sa* is most usual in the meaning ‘what?’, though.

Louisiana Creole

- (16) a. *Sa k’ale monje le chat?*
 IPW REL.NOM-is.going eat the cat
 ‘Who is going to eat the cat?’ (Neumann 1985:412 via Rottet 2004:183)
- a. *Ça qui tchué li?*
 IPW REL.NOM killed him
 ‘What killed him?’ (Wogan 1931:8 via Rottet 2004:183)
- b. *Ça to gaignin dan panier là?*
 IPW 2SG get in basket DEM
 ‘What do you have in the basket?’ (Wogan 1931:14 via Rottet 2004:183)

Non-human use of the interrogative *qui* is also attested in Laurentian French

¹⁶ The orthography of the examples in the source has been preserved. Note in this respect that *qui*/ki/ = *ki* and *ça*/sa/ = *sa*.

(Sudbury, Ontario, Canada), French of l'Île aux Coudres (St. Lawrence River, Québec, Canada), some Québécois varieties (Québec, Canada), French of Saint-Pierre and Miquelon (cf. Rottet 2004:180, 184). Finally, “*kisa (ki)* ‘what?’ is fairly widespread [in] French creoles of the Caribbean and the Indian Ocean” (Rottet 2004:184).¹⁷

By way of conclusion, a few words can be said on the origins of the non-human use of *qui*. Generally speaking, this appears to be due to the peculiarities of the way the system of interrogative and relative pronominals has been structurally reorganized and reduced in Old and Middle French. The distinction between the paradigms of the interrogative and relative pronominals was somewhat blurred already in Latin and the two paradigms fall completely together in French. The form M.SG.NOM *quī*, which functioned as both a relative and an interrogative, ousted the dedicated interrogative NON<N>.SG.NOM form *quis*. What is more, although in French the forms of the relative pronominals are identical to the forms of the interrogative pronominals, the two paradigms are organized on different bases. In the relative pronominals, the distinction is between nominative *qui* and accusative *que*, whereas in the interrogative pronominals the distinction is basically between human *qui* ‘who?’ and non-human *que* (and *quoi*) ‘what?’. This mismatch may explain why *qui* could have become ambiguous between ‘who?’ and ‘what?’ and particularly in the nominative. At the same time, the use of *qui* as ‘what?’ after prepositions seems to be due to a somewhat different reason. In all probability, in this function *qui* ‘what?’ results from the conflation of the nominative *qui* ‘who?’ and the reflex of the Latin dative form *cui*, which covered both the human and the non-human meanings. The reflex of the latter form, originally pronounced somewhat differently than *qui*, was used in French till approximately the 13th century as the common non-nominative form for both the human and the non-human interrogative pronominals (Bloch & von Wartburg 1968:525).

Finally, although in standard Middle and Classical French the distinction between human and non-human interrogatives seems to have been still maintained in the accusative, in a number of dialects the lack of differentiation between ‘who?’ and ‘what?’ became complete with only one general interrogative *qui* ‘who?, what?’ remaining. This feature was subsequently exported from these dialects into some colonial French varieties and creoles based on them.

¹⁷ Rottet (2004:184) does not say whether in any of these creoles *ki sa (ki)* can also be used as ‘who?’.

3.1.4 Old English

According to Karlberg (1954:7, 86-88, 96-98), Mitchell (1985:I, §§348-351), Mustanoja (1960:181-183) and the *Oxford English Dictionary* (2nd ed., entry for *what* A.I.2) among others,¹⁸ in Old English the interrogative pronominal *hwæt* ‘what?’ was regularly used as nominal predicate “in reference to name or identity [of a person], and thus equivalent to *who*” (*Oxford English Dictionary*, 2nd ed., entry for *what* A.I.2), as in (17), whereas the Old English interrogative *hwa* ‘who?’ was used in questions about persons and mostly in non-predicative functions, such as subject or object.

Old English

- (17) *þa cwæð Isaac: Hwæt eart þu?*
 then said Isaac **what** are you
He andwirde and cwæð: Ic eom Esau
 he answered and said I am Esau
 ‘Then said Isaac, “**Who** are you?” He answered and said, “I am Esau”’
 (Ælfric *Genesis* xxvii, ca. 1000, via *Oxford English Dictionary* 2nd ed., entry for *what* A.I.2; translation by Peter Petré, p.c.)

Mustanoja (1960:181-182) further claims that “there does not seem to be unambiguous cases of *who* as a predicate pronoun until about 1300” and that “in predicate position independent *what* is used for *who* from Old English down to the end of the 17th century” (cf. also similar statements in Karlberg 1954:96-98). However, the first claim appears to be an exaggeration, as can be illustrated by (18) (cf., e.g., Mitchell 1985:I, §§348-351 for further examples).

Old English

- (18) *hwa is ure nexta?*
 who is our closest.friend
 ‘Who is our closest friend?’ (Ælfric *Homilies* ii.318.1, ca. 1000, via Mitchell 1985:I, §§350; my translation)

Sweet (1898:§2119, via Mitchell 1985:I, §§350) notes that *hwæt* appears to be particularly preferred to *hwa* when the subject is a personal pronoun.

In Modern English predicative *what?* is used in questions about humans to inquire about their class membership, as in *What is he? A lawyer?*, their relation to other individuals, *What is he to you?*, etc., but not identity. Clear examples of such use are attested already in early Middle English, as in (19).

¹⁸ I am grateful to Peter Petré for suggesting the references.

Middle English

- (19) *whæt ært þu, fære wiht? eært þu angel, eært [þu] cniht[t]?*

what are you fair thing are you angel are you boy
 ‘What are you, fair maiden? Are you an angel, are you a boy?’ (*Lamon's Brut, or Chronicle of Britain*, ca. 1205, via *Oxford English Dictionary* 2nd ed., entry for *what* A.I.2; translation by Peter Petré, p.c.)

Arguably, this use of *hwæt* was also possible in Old English (20-22), but clear examples are rare.

Old English

- (20) *hwæt sind ða strangan, hwæt ða unstrangan?*
 what are the.PL strong.PL what the.PL un-strong.PL

[*Ʒa beoð strange and trume, ðe þurh geleafan and godum geearnungum wel ðeonde beoð. Ʒa sind unstrange þe...*]

‘Who/what are the strong ones, who/what are the weak ones? [Those are strong and firm who are prospering much through faith and good merits. Those are weak who...]’ (*Ælfric Homilies* ii.390.22, ca. 1000, via Mitchell 1985:I, §§351; my translation)

- (21) [*pam deofle wæs micel twynung,*] *hwæt Crist wære...*
what Christ could.be

[*þa smeade se deofol*] *hwæt he wære*
 what he could.be

[*hwæðer he wære Godes sunu, seðe manncynne behaten wæs*]

‘[Devil was in great doubt (about)] what/who Christ was... [and the devil thought (about)] what/who he was; [whether he was a son of God, the one that was promised to mankind]’ (Ælfric *Homilies* i.14.4, ca. 1000, via Mitchell 1985:I, §§351; my translation)

- (22) *he nyste hwæt hie wæron*
 he did.not.know what they were

‘He did not know what they were’ (“‘He did not know who they were’, that is, ‘he did not know what their nationality was’”, Mitchell 1985:I, §§350)

The presumed predicative use of *hwæt* ‘what?’ instead of *hwa* ‘who?’ in questions about the identity of a person is rather strange, particularly because *hwæt* is originally the neuter form of the same interrogative pronominal of which *hwa* is the masculine form. Taking into account the fact that in general, in Indo-European languages there are many masculine non-human nouns and only few neuter human nouns, it would have been more natural if the masculine *hwa* had

expanded into the domain of the neuter *hwæt* rather than the other way round. Of course, it is possible that the Old English predicative *hwæt* in questions about humans does not inquire about the identity of humans but rather is used in the same way as the Modern English *what?* in *What is he?* This would mean that all the examples adduced in the sources have been misinterpreted. However, at least in examples such as (17) above, this appears to be hardly plausible. Therefore, it seems that indeed, next to its regular use as ‘what?’ in both predicative and non-predicative functions, the Old English predicative *hwæt* was also used to inquire the identity of humans, as well as about their classification.

From a cross-linguistic perspective, Old English may be rather special, but it is definitely not unique. Thus, in many respects, the distribution between *hwæt* and *hwa* in Old English appears to be similar to the distribution between *ma* and *mi* in Ahaggar Tuareg (Afro-Asiatic, Berber; Algeria), as discussed in Section III.2.3.2.2.2. Compare also Bijogo (Niger-Congo, Atlantic; Guinea-Bissau) in Section III.2.1.2.

As far as the origins of the predicative use of *hwæt* as ‘who?’ are concerned, it seems reasonable suppose that this use developed out of the predicative use of *hwæt* in questions involving a non-prototypical combination of values of the kind [person + classification (+ common noun)], i.e. KIND-questions (cf. Section II.2). Such a shift from classification to identification with predicatively used interrogatives would have parallels in other branches of Indo-European. Thus, recall the predicative use of French *quel* (basically an attributive interrogative ‘which [N]?, what (kind of) [N]?’) as ‘what?’ and ‘who?’ and earlier also in KIND-questions (cf. Section III.3.1.3.1). Consider also the development of Ancient Greek *poios* ‘what (kind of) [N]? (sometimes also which [N]?)’ into Modern Greek *poios* ‘who?’ (and attributive ‘what (kind of) [N]?, which [N]?’), which presumably occurred through its predicative use in KIND-questions (cf. Section III.3.1.5).¹⁹ However, unlike in the latter two cases, by Old English

¹⁹ In fact, Old English *hwæt* (originally neuter) and *hwa* (originally the masculine counterpart of *hwæt*) also go back to the same Indo-European interrogative stem **k^wo-/*k^we-* (with feminine **k^wā-*) as Greek *poios*, French *quel* (from Latin *quālis*) or Latin *quī* (from earlier *quoi*) and *quod* rather than Indo-European **k^wi-*, as Greek *tis* and *ti* and Latin *quis* and *quid*. Note, in this respect, that originally the Latin and Greek interrogatives based on **k^wo-/*k^we-* appear to have primarily functioned attributively as ‘what (kind of) [N]?’ (sometimes also ‘which [N]?’), whereas the interrogatives based on **k^wi-* were mostly substantival in both languages (cf. Sections III.3.1.5-III.3.1.6). Given this and the fact that several adjectival suffixes of the form **(C)o*, such as **-jo*, have been reconstructed for Proto Indo-European (cf., e.g., Watkins 1998:63-67), it seems reasonable to consider the possibility that in origin, **k^wo-/*k^we-* may be just some adjectival derivations of **k^wi-* and that they were primarily used to inquire about quality as ‘what kind of [N]?’. Traditionally, however, no semantic or functional

times, *hwæt* also functioned as a full-fledged ‘what?’ interrogative in both predicative and non-predicative uses. This conflict between the two meanings of *hwæt*, ‘who?’ in predicative use and ‘what?’ in both predicative non-predicative uses, was later resolved in favour of ‘what?’.

3.1.5 Ancient Greek

Strictly speaking, Ancient Greek does not have interrogative pronominal forms exclusively dedicated to expressing the opposition human vs. non-human, such as ‘who?’ vs. ‘what?’, or ‘which one (human)?’ vs. ‘which one (non-human)?’. Instead, it distinguishes between a non-neuter (which in fact is masculine in both its origin and form) interrogative pronominal *tis* (NON<N>.SG.NOM) and a neuter one *ti* (N.SG.NOM), on the one hand, and an interrogative *poios* (M.SG.NOM), *poia* (F.SG.NOM), *poion* (N.SG.NOM), on the other hand.^{20, 21}

The latter is usually attributive ‘what (kind of) [N]?’ (sometimes also ‘which [N]?’), but can also be used independently as a predicate in the meaning ‘what?, of what kind?’ (sometimes, also ‘which one?’), also in questions about humans, as in (23).

Ancient Greek

- (23) *poiōi* *k’* *eit’* *Odusēi*
 IPW.M.PL.NOM PTCL be.PRS.OPT.ACT.2PL PROP.M.DAT.SG
amunemen
 keep.off.PRS.OPT.INF.ACT
 ‘What are you to keep off Odusseus?’ (Liddell *et al.* 1940 under *poios*)

Interestingly, in Modern Greek, although *poios* M.SG (*poia* F.SG, *poio* N.SG) is still used attributively, the masculine form *poios* (and marginally the feminine *poia*) is also the interrogative pronominal used (both predicatively and non-predicatively) to inquire about the identity of humans, i.e. ‘who?’, and not their classification. The development of *poios* from ‘what (kind of) [N]?’ into ‘who?’ has probably occurred through its predicative use in questions involving a non-prototypical combination of values of the kind [person + classification (+ common noun)], i.e. KIND-questions. In particular, *poios* must have been first

difference is reconstructed between the two stems.

²⁰ Most diacritical marks are usually left out in transliterations of Ancient Greek. As a rule, I will not mark the diacritical signs either. The transliterations *ê* and *ô* in the examples stand for the letters *η* and *ω*, respectively.

²¹ In fact, there are a couple of other interrogatives, such as *póteros* (M.SG.NOM) ‘which (one) of the two?’ and *potapos* (M.SG.NOM) ‘from what country?, where from?, what (kind of) [N]?’ but they are less relevant here.

expanded in its predicative use to questions about the identity of persons, thus becoming rather similar to the French interrogative *quel* (cf. Section III.3.1.3.1). In questions about things, though, the neuter counterpart of *poios*, viz. *poio*, must have lost the competition to the interrogative pronominal *ti* ‘what?’ (originally the neuter counterpart of *tis*). As a result, *poios* must have become restricted to questions about persons, either their identity or classification. Subsequently, it was ousted by *ti* ‘what?’ from KIND-questions as well (again compare French *quel* in this respect). Finally, in Modern Greek, we find it generalized as ‘who?’ to all kinds of syntactic contexts, both predicative and non-predicative.

The interrogative *tis* is inflected for gender, number and case, as illustrated in Table 2. I have highlighted with grey shades the instances of neutralization of gender distinctions within this paradigm.

Table 2. Paradigm of the Ancient Greek interrogative *tis* (Kühner 1890-1904:I, §176)

	SG		PL		DU
	M (& F)	N	M (& F)	N	M (& F), N
NOM	<i>tis</i>	<i>ti</i>	<i>tines</i>	<i>tina</i>	<i>tine</i>
ACC	<i>tina</i>		<i>tinas</i>		
GEN	<i>tinós</i>		<i>tinôn</i>		<i>tinoin</i>
DAT	<i>tini</i>		<i>tisi(n)</i>		

As can be readily observed, in the dual there is no distinction in gender, which suggests that in the dual no distinction can be made between the meanings ‘who?’ and ‘what?’. However, the dual is very rare in texts and most of the few examples of *tine* I found, such as (24), are questions about things.

Ancient Greek

(24) *tine toutô?*

IPW.DU.NOM DEM.DU.NOM

‘[A: In the case of the body there are two arts which have to do with these two evil conditions, are there not? B:] What are they?’ (Plato *Sophist* 228e at <http://www.perseus.org>, reference via Marc Huys, p.c.)

Example (25) is a question about persons.

Ancient Greek

- (25) *sphô tin' eston?*
 2DU.NOM IPW.DU.NOM be.PRS.ACT.IND.2DU
nô, brotô
 1DU.NOM mortal.man.DU.NOM
 '[Epops (a bird): I am a bird [...]] All birds moult their feathers, you know, every winter, and others grow in their place. But tell me,] who are you (two)? [Euelpides:] We (two)? We (two) are (two) mortals' (Aristophanes *Birds* 107 at <http://www.perseus.org>, reference via Liddell *et al.* 1940 under *tis*)

In singular and plural, the distinction between the neuter and non-neuter forms of the interrogative pronominal is maintained in the nominative and accusative. At least in predicative use or in selective contexts, the interrogative pronominal should agree in gender with the subject. Therefore, one would expect that it is possible in such contexts to use both the non-neuter *tis* and neuter *ti* to inquire about both humans and non-humans, depending on gender. However, given that human nouns only rarely belong to the neuter gender (usually, diminutives), but there are many non-human nouns that belong to the masculine and feminine genders, it would be expected that *ti* should be hardly used in questions about humans, while *tis* should be found in questions about non-humans, as in (26).²²

Ancient Greek

- (26) *tis gar meizôn estin*
 IPW.M.SG.NOM because great.CMPR.M.SG.NOM be.PRS.ACT.IND.3SG
ho chrusos ê ho
 DEF.ART.M.SG.NOM gold.M.SG.NOM or DEF.ART.M.SG.NOM
naos?
 temple.M.SG.NOM
 '[Ye fools and blind:] for whether is greater, the gift, or the altar [that sanctifieth the gift]?' (Matthew 23:17 in the Septuagint as available at <http://www.zhubert.com>)

The interrogative forms *tis* and *ti* are also common in attributive use. In this case, the choice between the two is determined by the gender of the noun they modify, masculine/feminine vs. neuter, as in (27).

²² A better example is needed, since the context here may also be interpreted as selective.

Ancient Greek

- (27) a. *tines* *aneres*
 IPW.M.PL.NOM man.M.PL.NOM
 ‘which/what men?’ (Bailly 1901:873)
- b. *tis* *achô*
 IPW.F.SG.NOM sound.F.SG.NOM
 ‘which/what sound?’ (Liddell *et al.* 1940 under *tis*)
- c. *ti* *perisson* *poieite*
 IPW.N.SG.ACC prodigious.N.ACC do.PRS.ACT.IND.2PL
 ‘[If you greet only your brothers,] what more (lit.: ‘what prodigious’) are you doing [than others? Do not even the Gentiles do the same?]
 (Matthew 5:47 in the Septuagint as available at <http://www.zhubert.com>)

3.1.6 Latin

Latin appears to be quite similar to Ancient Greek. Thus, strictly speaking, Latin does not have interrogative pronominal forms exclusively dedicated to expressing the opposition human vs. non-human, such as ‘who?’ vs. ‘what?’, or ‘which one (human)?’ vs. ‘which one (non-human)?’. Instead, it has two masculine interrogatives, *quis* and *quī*, which are inflected for number and case and have neuter and feminine counterparts. As illustrated in Table 3, their paradigms almost completely overlap. In many cases, gender distinctions are also neutralized. For the sake of clarity, I highlighted such forms with grey shades in Table 3.

Table 3. Gender-number-case forms of the Latin interrogative pronominals (e.g., Allen & Greenough 1903:I.§148; Lavency 1985:§120, §§123-124; Madvig 1870:§86, §88)

	SG			PL		
	M	N	F	M	N	F
NOM	<i>quis, quī</i>	<i>quid, quod</i>	<i>quae</i>	<i>quī</i>		<i>quae</i>
ACC	<i>quem</i>	<i>quid, quod</i>	<i>quam</i>	<i>quōs</i>		<i>quās</i>
GEN	<i>cūius</i>			<i>quōrum</i>		<i>quārum</i>
DAT	<i>cui</i>			<i>quibus</i>		
ABL	<i>quō</i>		<i>quā</i>			

All the forms except *quis* and *quid* can also be used as relative pronouns. As interrogatives the neuter singular forms *quid* and *quod* differ in that *quid* only

occurs in substantival functions, while *quod* is used attributively. The difference between the masculine singular forms *quis* and *quī* (from earlier *quo-i*) is, generally speaking, of the same kind but it is much less strict. Thus, although *quis* is normally used in substantival functions, it can also be used attributively, as in *quis diēs fuit?* ‘what day (M.SG.NOM) was it?’ or *quis homō* ‘what man (M.SG.NOM)?’ (Allen & Greenough 1903:I.§148). Similarly, although *quī* is normally attributive, as in *quī homō* ‘what kind of man?’, it can also be used as a nominal, as in *quī nōminat mē?* ‘who calls my name? (lit.: ‘who names me?’)’ (Allen & Greenough 1903:I.§148). However, according to Kühner & Holzweissig (1974:I.§141), unlike *quis*, a nominal *quī* is often classificational ‘what kind of man?’ rather than purely identificational ‘who?’, as illustrated in (28).

Latin (Kühner & Holzweissig 1974:I.§141)

- (28) a. *quis* *me* *vocat?*
 IPW.M.SG.NOM 1SG.ACC call.PRS.ACT.IND.3SG
 ‘Who calls me?’
- b. *quī* *me* *vocat?*
 IPW.M.SG.NOM 1SG.ACC call.PRS.ACT.IND.3SG
 ‘What kind of man calls me?’

The feminine forms seem to be rare in substantival functions. In other words, *quis*, and to a lesser extent *quī*, are the default interrogatives used in questions about humans and, thus, the closest equivalents of ‘who?’, while *quid* is the default interrogative used in questions about non-humans and, thus, the closest equivalent of ‘what?’. However, just as in Ancient Greek, Latin interrogatives are expected to be used predicatively and in this case they should normally agree with the subject in gender. Like in Ancient Greek, in Latin there are many non-human nouns belonging to the masculine and feminine genders, which entails that the interrogative used as a nominal predicate with such a non-human noun as subject should agree with the latter in gender, as illustrated in (29) with masculine non-human nouns and in (30) with feminine non-human nouns.²³

Latin

- (29) a. *qui(s)* *ille* *piscis* *est?*
 IPW.M.SG.NOM DEM.M.SG.NOM fish.M.SG.NOM is
 ‘What is that fish?’

²³ According to Marc Van Uytanghe, Dirk Sacré and Pierre Swiggers (p.c.), these constructed examples represent correct Classical Latin.

- b. *qui(s)* *ille* *fluvius* *est?*
 IPW.M.SG.NOM DEM.M.SG.NOM river.M.SG.NOM is
 ‘What is that river?’
- (30) a. *quae* *illa* *res* *est?*
 IPW.F.SG.NOM DEM.F.SG.NOM thing.F.SG.NOM is
 ‘What is that thing?’
- b. *quae* *illa* *urbs* *est?*
 IPW.F.SG.NOM DEM.F.SG.NOM town.F.SG.NOM is
 ‘What is that town?’

3.2 North Caucasian languages

North Caucasian languages are spoken chiefly in the north Caucasus regions of Russia and neighbouring Georgia and Azerbaijan. Some small groups of the speakers of North Caucasian languages are also found in Turkey, but they migrated there only in the second half of the 19th century. North Caucasian languages are subdivided in two major branches, East Caucasian (or Nakh-Dagestanian) and West-Caucasian (or Abkhaz-Adyghean).

Languages with interrogative pronominal wordforms that can be used indiscriminately as ‘who?’ or ‘what?’ in all syntactic functions appear to exist both in the East and West Caucasian branches. However, they are not numerous, in all probability just three in total: Kidero Tsez (East Caucasian), Tapanta Abaza (West Caucasian), and Abkhaz (West Caucasian). The lack of differentiation between ‘who?’ and ‘what?’ in all the three languages is mostly due to the peculiarities of the gender-number morphology, which distinguishes (or used to distinguish) between human and non-human meanings in the interrogative forms at issue. Furthermore, in some varieties of Tsez and a closely related language Inkhokari Khvarshi, the lack of differentiation between ‘who?’ and ‘what?’ can be found only in the predicative function, while in the non-predicative function the distinction between ‘who?’ and ‘what?’ is partially neutralized in case. In what follows, first the East Caucasian branch will be discussed (Section III.3.2.1) and then the West Caucasian (Section III.3.2.2).

3.2.1 East Caucasian

East Caucasian languages are primarily spoken in the Chechen and Dagestan Republics of the Russian Federation and neighbouring regions of Georgia and Azerbaijan. East Caucasian languages are also often referred to as Nakh-Dagestanian, following the presumed subdivision of these languages into two major branches, Nakh languages and Dagestanian languages. In other classifications, Nakh languages are considered to be on a par with the language

groups otherwise considered together as Dagestanian: Avar-Andic, Tsezic, Dargi, Lak, Khinalug, and Lezgi.²⁴ Tsezic languages are usually grouped together with Avar-Andic languages as Avar-Andic-Tsezic languages, while Dargi is often grouped with Lak, and Khinalug with Lezgi.²⁵

There appears to be only one East Caucasian language with an interrogative pronominal that can be applied indiscriminately as ‘who?’ or ‘what?’ in all syntactic functions, viz. the Kidero dialect of Tsez (or Dido), as described by Bokarev (1967b). Together with Hinukh and Khvarshi, Tsez forms the West Tsezic group, the other two Tsezic languages, Hunzib and Bezhta, constitute the East Tsezic group. Tsez is spoken by ten to twelve thousand people in some 50 villages concentrated primarily in the Tsuntin District in the west of Dagestan (cf. Khalilov 2005:346-347). From 2 to 5 major dialects are distinguished, with many lower level variants. The variety of Tsez described in Bokarev (1959, 1967b) is spoken in the village Kidero. The idiom spoken in this and several neighbouring villages (Gutatl’, Zakhida, Azil’ta, and Mokok) is usually referred to as the Kidero dialect of Tsez. It should be noted, however, that the idioms spoken in the latter villages differ to some extent from that of Kidero, in particular the idioms spoken in Azil’ta and Mokok (Bokarev 1967b:418).

The lack of differentiation between ‘who?’ and ‘what?’ in Kidero Tsez is primarily due to the peculiarities of the development of its gender-number agreement morphology, originally used on the interrogative pronominals. In what follows, I will first present the Tsez interrogative pronominals. Then, the origins of the lack of the differentiation between ‘who?’ and ‘what?’ in Tsez will be discussed in more detail.

The system of the interrogative pronominals of Kidero Tsez, as described by Bokarev (1967b), is reproduced in Table 4.

Table 4. Kidero Tsez interrogative pronominals (Bokarev 1959:199, 1967b:411)²⁶

	G1	G2, G3, G4
Absolute		<i>šeb(i)</i>
Ergative	<i>ʃu</i>	<i>ʃina</i>
Other cases	<i>ʃu-</i> + CASE	<i>ʃina-</i> + CASE

As can be seen in Table 4, instead of the human vs. non-human distinction, the

²⁴ For instance, this classification is used by the Ethnologue.

²⁵ See, for instance, Alekseev (2001c, 2001d) for a brief overview of the extant classifications of the East Caucasian languages.

²⁶ The interrogative pronominals do not have separate plural forms.

interrogative pronominals distinguish gender. Gender assignment in Tsez is mostly semantic, at least for animates, including humans. The way it correlates with a human vs. non-human and animate vs. inanimate distinction is summarized in (31).

(31) *Gender assignment in Kidero Tsez (based on Bokarev 1967b:407)*

- G1 male humans
- G2 female humans and some inanimates
- G3 animates
- G4 inanimates

As can be readily observed, this distribution of human nouns among the two genders, G1 and G2, implies that when a question is based on a presupposition that the person in issue must be a woman²⁷ the same form of the interrogative will be used as in questions about non-humans. Unfortunately, Bokarev does not provide examples of such contexts. Examples (32) and (33) illustrate the use of the absolutive form *šeb* for both humans and non-humans.

Kidero Tsez (Bokarev 1959:200)

- (32) *mi šeb esir-χo?*
 2SG.ERG IPW.ABS ask.for-PRS
 ‘What are you asking for?’
- (33) *šeb Ø-aj-a:?*
 IPW.ABS AG1-come-[PST]Q
 ‘Who has come?’²⁸

As can be deduced from (33) and (34), the G1 form of the interrogative must be the unmarked form in questions presupposing a human referent of whom the sex is unknown.

²⁷ For instance, as in *Whose bra is this?* or, given that clothes are traditionally made by women (cf. Rajabov 1997:128), *Who made this beautiful dress?*, etc.

²⁸ Note that in Tsez only a small minority of verbs is marked for gender agreement, including *aj-* ‘come’, but not *esir-* ‘ask for’ (cf. Bokarev 1959:205-206). This explains why the zero AG1 marker *Ø-* is present in (33) and no marker at all in (32). The agreement pattern of G1 is expressed with a zero marker, while the agreement patterns of other genders have overt markers. It should also be mentioned that in the source itself the verb was written *aja*, but *aj-a* is the so-called infinitive I form of the verb ‘come’ (cf. Bokarev 1959:211 or 1967b:413), which is both impossible here and incompatible with the translation provided. The final vowel of the verb should probably be long, as this vowel is in all probability the interrogative suffix *-a:*, which in the so-called past I is added directly to the verbal base (cf. Bokarev 1959:217). In this respect, compare the verb form *etŋ-a:* in (34).

Kidero Tsez (Bokarev 1959:200)

- (34) *ɬu* *eɬa:* *debzo* *esjuqχor* *žo*
 IPW.ERG[AG1] say-[PST]Q 2SG.GEN brother[G1] that.SG.ABS
χabar?
 news[SG.?G4]
 ‘Who told your brother about that news?’

It should be mentioned that not all sources describe the Tsez interrogative pronominals in the same way as Bokarev (1967b), which is most likely due to a difference in the dialects involved. Consider, for instance, the interrogative pronominals of the Tsebari subdialect of the Asakh dialect of Tsez summarized in Table 5.

Table 5. Interrogative pronominals of the Tsebari subdialect of the Asakh dialect of Tsez (http://en.wikipedia.org/wiki/Tsez_language)

	‘who?’	‘what?’
Absolute		<i>šebi</i>
Ergative	<i>ɬu</i>	<i>ɬina</i>
Other cases	<i>ɬa:-</i> + CASE	<i>ɬina-</i> + CASE

Here, the distinction between ‘who?’ and ‘what?’ is neutralized in the absolute case. However, since the absolute is also the case of nominals in the predicative function, we can say that this variety of Tsez is characterized by a lack of differentiation between ‘who?’ and ‘what?’ in the predicative function, whereas in the non-predicative function the distinction between ‘who?’ and ‘what?’ is partially neutralized in case.

Khalilov (2001) describes yet another system with fully differentiated ‘who?’ and ‘what?’, Table 6. He does not specify the dialect described.

Table 6. Tsez interrogative pronominals according to Khalilov (2001:327)

	‘who?’	‘what?’
Absolute	<i>šow</i>	<i>šebi</i>
Ergative	<i>ɬu</i>	<i>ɬina</i>
Other cases	<i>ɬu-</i> + CASE	<i>ɬina-</i> + CASE

Let us now discuss the origins of the lack of differentiation between ‘who?’ and ‘what?’ in Kidero Tsez in more detail. This lack of differentiation appears to be primarily due to the peculiarities of the development of the gender-number agreement morphology originally used on the interrogative pronominals. Thus,

the final *-w* and *-b(i)* in the absolutive forms of the Tsez interrogative pronominals (and perhaps also *-u-* and *-i-* in the non-absolutive forms) seem to be frozen reflexes of the gender agreement markers, which in Proto Tsezic and in all probability also in Proto East Caucasian, used to distinguish the human interrogative pronominals, masculine and feminine, on the one hand, from the non-human interrogative ones, on the other. Later, the system of these markers on the interrogative pronominals was reorganized in Kidero Tsez in such a way that the original distinction was first transformed into male human (AG1) vs. the rest and then was completely lost in the absolutive case. Interestingly, the demonstratives and anaphoric pronominals of Kidero Tsez have undergone a similar development, as illustrated in Table 7 on the example of the singular distal visible demonstrative *(how)žo* also used as an anaphoric pronominal. Compare Table 7 to Table 4.

Table 7. Kidero Tsez singular anaphoric and distal visible demonstrative pronominal (Bokarev 1967b:410-411)

	G1	G2, G3, G4
Absolutive	<i>(how)žo</i>	
Ergative	<i>nesa</i>	<i>neła</i>
Other cases	<i>nesi-</i> + CASE	<i>neło-</i> + CASE

Similar conflations of the human and non-human agreement patterns within the gender-number paradigms of demonstratives-cum-anaphors are common in Tsezic languages (e.g., see Bokarev 1967c:445 on Hinukh or Bokarev 1967d:479-480 on Hunzib). At the same time, this is not so common in the interrogative pronominals, where usually the conflation takes place between the two human agreement patterns, AG1 for male humans and AG2 for female humans, on the one hand, and all other non-human patterns, on the other, just like in the Tsez variety described by Khalilov (2001) (cf. Table 6 above). In this respect, Kidero Tsez is exceptional indeed, although some other dialects of Tsez, such as the Tsebari subdialect of Asakh Tsez presented in Table 5 above, and Inkhokari Khvarshi, Table 8, come rather close to Kidero Tsez in neutralizing the distinction between a human and a non-human interrogative pronominal in the absolutive case, and consequently, as was pointed out with respect to Tsebari Tsez above, lacking differentiation between ‘who?’ and ‘what?’ in the predicatively used interrogative pronominals.

Table 8. Interrogative pronominals of Inkhokari Khvarshi (Bokarev 1967a:428)

	‘who?’	‘what?’
Absolutive		<i>hibo</i>
Ergative	<i>ɬu</i>	<i>ɬini</i>
Genitive I	<i>ɬijo</i>	<i>ɬine-</i> + CASE
Other cases	<i>ɬu-</i> + CASE	

The attested conflation between the human and non-human agreement patterns within the gender-number paradigms of demonstratives-cum-anaphors and interrogatives in Tsezic languages is due to a concurrence of various semantic and formal changes the foundation for which must have already been laid in Proto Avar-Andic-Tsezic times and in all probability even much earlier. Given that a detailed reconstruction of these changes falls outside of the scope of the present study, let us just briefly consider the gender-number systems reconstructed, for instance, for Proto Tsezic and Proto East Caucasian, paying particular attention to those features of these systems that may have been relevant in the processes at issue.

In Tsezic, and in East Caucasian languages in general, gender is not marked on the nouns themselves. Instead, gender manifests itself through gender agreement markers on verbs, adjectives, demonstratives, anaphoric pronominals and interrogative pronominals. Typically, the agreement markers appear as prefixes on verbs and adjectives and as suffixes on demonstratives, anaphors and interrogatives. The gender system reconstructed for Proto Tsezic is reproduced in Table 9 and the system usually reconstructed for Proto East Caucasian is presented in Table 10. The reconstructed semantics of the genders is provided in brackets.

Table 9. The Proto Tsez gender system (Alekseev 2001a:301)

	G1 (male humans)	G2 (female humans)	G3 (animates & some inanimates)	G4 (inanimates)	G5 (inanimates)
SG	* <i>w</i>	* <i>j</i>	* <i>b</i>	* <i>j</i>	* <i>r</i>
PL	* <i>b</i>	* <i>b</i>	* <i>b</i> ²⁹	* <i>r</i>	* <i>r</i>

²⁹ In an earlier version of this reconstruction (Alekseev 1988:154), the plural agreement marker of gender G3 was reconstructed as *-*r*.

Table 10. The Proto East Caucasian gender system (based on Dumézil 1933:29; Alekseev 1988:184-189, 2001d:164-165)

	G1 (male humans)	G2 (female humans)	G3 (animates & some inanimates)	G4 (inanimates)
SG	* <i>w</i>	* <i>j</i>	* <i>b</i>	* <i>r</i>
PL	* <i>b</i>	* <i>b</i>	* <i>r</i>	* <i>r</i>

As can be readily observed, the number of different forms of the gender-number agreement markers is considerably smaller than the number of cells in the two tables. Furthermore, the human and non-human gender-number agreement markers turn out to be very similar. Consider, for instance, the Proto East Caucasian plural human agreement marker AG1~AG2 **b* and the singular non-human agreement marker AG3 **b*, or the Proto East Caucasian singular human male agreement marker AG1 **w* and the singular non-human agreement marker AG3 **b*. The other three agreement markers, **j* and two **r*-s, are also formally not so divergent, given that, for instance, the shift **r* > *j* is attested in some Tsezic languages (cf. Alekseev 1988:186). This original strong formal resemblance between the human and non-human gender-number agreement markers but not between the two singular human gender-number agreement markers provides a particularly fertile ground for the development of such situations when the same form functions as a human gender-number marker in one daughter language but as a non-human gender-number marker in the other.³⁰ Note also that as has been mentioned above, most typically gender-number agreement markers become neutralized on demonstratives-cum-anaphors and interrogative pronominals, rather than verbs or adjectives. In all probability, this

³⁰ Consider in this respect a remark made by Talibov (1976:72, cited via Alekseev 1988:184; translation and comments are mine): “Although all Dagestanian languages use the same set of forms to mark gender-number agreement, a given form may be found in several absolutely different functions. Thus, if a given form [i.e., *w*] is used as a gender agreement marker in Avar-Andic languages with [singular] gender 1 nouns [i.e., male human nouns], the same marker in some Lezgian languages is used with [singular] gender 3 nouns [i.e., non-human animate and inanimate nouns; cf., e.g. Alekseev (2001e:371) on Lezgian in general and Alekseev (2001b:411-412) on Rutul in particular], while in Lak it is used either with [singular] gender 1 nouns [i.e., male human nouns] or [singular] gender 3 nouns [i.e., most female human nouns, as well as most non-human animates and many inanimate nouns (Murkelinsky 1967:491)].” Alekseev (1988:184) comments on this quote by noting that “most likely, there is no need here to look for an explanation of these discrepancies anywhere else than in the patterns of phonological change”.

is due to the fact that these markers are suffixes on the former elements, but prefixes on the latter.

Besides the formal factors favouring conflation of human and non-human gender-number agreement markers, there are certain peculiarities of gender assignment principles and of the semantics of the agreement patterns of Avar-Andic-Tsezic languages that must have further corroborated the blurring of the strict division between human and non-human genders and the respective agreement patterns. Thus, it is not uncommon for gender G2, which is presumed to be originally an exclusively female human gender, to include various non-human nouns. For instance, according to Rajabov (1997:128), in Tsez many nouns designating inanimate entities are assigned to this gender because in one way or another, they are traditionally related to women, e.g. many names of clothes, because they are traditionally made by women, names of berries, because they are gathered by women, the words for milk, although normally liquids are assigned to G3. In some languages, according to Alekseev (1988:104, 187), the singular AG2 marker **j* may be used to express the diminutive meaning, as in (35b) vs. (35a) from Tindi, irrespective of whether they denote a female human or not.

Tindi (Avar-Andic; Russia; Gudava 1967b:371)

- (35) a. *bu-k^w-o* *ce-b* *sari*
 AG3.SG-be-PST one-AG3.SG fox[G3.SG]
 ‘There was one fox.’
- b. *ji-k^w-o* *ce-j* *sari*
 AG2.SG-be-PST one-AG2.SG fox
 ‘There was one small fox.’

Having discussed the peculiarities of the development of the gender-number agreement morphology originally used on the interrogative pronominals, by way of conclusion a few words can be said on the interrogative pronominal roots themselves.

First of all, the use of a suppletive form for the non-absolutive cases of the interrogative pronominals, as in Kidero Tsez (cf. Table 4 above), is typical of the East Caucasian languages in general (cf. Alekseev 2001d:159). This feature squares well with a similar use of a special non-absolutive base in nouns, often coinciding with the ergative (cf. Alekseev 2001d:158). Normally, this base is not suppletive, though. Instead, it is derived from the absolutive base through the addition of some (frozen) suffixal morphology. In all probability, the element *ɬ*- in the non-absolutive forms of the Kidero Tsez interrogative pronominals is also a reflex of a similar non-absolutive base suffix with a loss of the original interrogative root to which this formative used to be added. The element *ɬ*- resembles various East Caucasian locative case markers (e.g., cf. Alekseev

1985:24, 51-52, 1988:85-86). A similar development is for instance suggested by Alekseev (1985:74) for the Lezgi ergative *ni* ‘who.ERG?’, which he compares to the Richa Agul *na* ‘who.ERG?’ vs. the Tpig Agul *hi-na* ‘who.ERG?’, where *hi-* is a reflex of the Proto Lezgitic interrogative **hi* ‘which one?’.

The element *še-* in the absolutive form of the Kidero Tsez interrogative pronominals is clearly related to the Proto Avar-Andic interrogative pronominal base **š:i-* + -AG ‘who?, what?’ (cf. Alekseev 1988:134) and to the similar Proto Lezgitic interrogative pronominal base **ši-* (cf. Alekseev 1985:74). Interestingly, according to Alekseev (1985:84), the latter interrogative must have been used only predicatively, i.e. in constructions like *What/who is this?*, which squares well with the fact that as such it is indifferent to the opposition human vs. non-human. The two interrogative stems **hi-* and **š(:)i-* may have been differentiated as selective vs. non-selective respectively.

3.2.2 West Caucasian

West Caucasian languages are primarily spoken in the Kabardino-Balkarskaya, Karachaevo-Cherkesskaya and Adygeya Republics of the Russian Federation, and in the Abkhazia Republic of Georgia. West Caucasian languages are also often referred to as Abkhaz-Adyghean. The received view is that West Caucasian consists of three branches: Abkhaz-Abaza, Circassian (including, Adyghe and Kabardian), and Ubykh.

There are two West Caucasian languages with a word that can be used indiscriminately as ‘who?’ or ‘what?’, the Tapanta dialect of Abaza, as described by Genko (1955), and Abkhaz. Roughly speaking, the lack of differentiation between ‘who?’ and ‘what?’ in both languages is due to the neutralization of the human vs. non-human distinction in the plural. In what follows, I will first discuss the data from Tapanta (Section III.3.2.2.1) and then the standard Abkhaz data (Section III.3.2.2.2).

3.2.2.1 Tapanta Abaza

Abaza is spoken by some thirty thousand people, primarily in the north of the Karachaevo-Cherkesskaya Republic. It has two major dialects, Ashkaraua (or Shkaraua) and Tapanta. Tapanta has formed the basis of the standard language.

Before proceeding, a caveat should be made. Tapanta Abaza is a highly synthetic language that strictly speaking does not have interrogative pronominals. Instead, it normally uses an interrogative bound root (or roots, depending on the analysis), which is indifferent to the human vs. non-human distinction. As such, this is interesting but of little immediate relevance for us because we are looking for ‘who?, what?’ words rather than just roots. The closest parallel to the interrogative pronominal words of languages like English that one may find in

Tapanta Abaza is provided by some nominalized verbal wordforms, such as *dazəwda* (or contracted *dəzda*) ‘who (exactly)? (SG)’ *dzač^wəja* ‘who? (SG)’, *(j)zač^wəja* (rarely *ač^wəja*) ‘what? (SG)’ or *zač^wək^wəja* ‘who?, what? (PL)’ (Genko 1955:105; Lomtatidze 1967b:134). As can be seen, the plural form is a full-fledged word and it is indifferent to the human vs. non-human distinction, which is exactly what we are looking for. In this respect, consider also *dzapšəja* ‘like who?, resembling who? (SG)’ and *(jə)zapšəja* ‘like what?, resembling what? (SG)’ vs. *jəzapšk^wəja* ‘like/resembling who?, like/resembling what? (PL)’ (Genko 1955:106).

In accordance with what has just been said, in what follows I will first present the Tapanta Abaza interrogative bound root(s). Then, I will discuss the nominalized verbal wordforms used in this language as independent interrogative pronominals, paying particular attention to the reasons accounting for the lack of differentiation between a human and a non-human meaning in the plural forms.

The bound interrogative morphemes used by Tapanta Abaza to form constituent questions are *-j-a* ‘what?’ (36a, 37a, 38a) and *-d-a* ‘who?’ (36b, 37b, 38b). Note that the two sequences can be further subdivided into a third person singular intransitive subject marker, non-human *j(ə)-* and human *d(ə)-*, and a general interrogative root *-a*, which probably should also be seen in the adverbial interrogative base *-b-a* (as in *a-ba* ‘where (at, to, from)?’, *an-ba* ‘when?’, *š-pa* ‘how?’, *aχ^j-pa* ‘whither?’).³¹

Tapanta Abaza

- (36) a. *wə-z-zə-pšəa-wa-j-a?*
 2SG.M[S]-PTCP.OBL-APPL-look-PTCP.IPFV-3SG.NON<HUM>.S-IPW
 ‘What are you looking for? (lit.: ‘The one that you are looking for, it is who/what?’)’ (Genko 1955:106)
- b. *wə-z-zə-pšəa-wa-d-a?*
 2SG.M[S]-PTCP.OBL-APPL-look-PTCP.IPFV-3SG.HUM.S-IPW
 ‘Who are you looking for? (lit.: ‘The one that you are looking for, s/he is who/what?’)’ (Genko 1955:106)

³¹ In all probability, this element *b* is related to the nominal predicate suffix *-b*, as in *a-rə-j mza-b* ‘this is a lamp’ (lit.: DEF-this-SG lamp-COP) (Genko 1955:71), and the so-called “deictic particle” *-ba-* used to derive emphatic demonstrative forms from the full forms of ordinary demonstratives, e.g. the short form *a-ra* ‘this (close to me)’ (lit.: DEF-this) ‘this’ → the full form *a-rə-j* ‘this’ (lit.: DEF-this-SG) → the emphatic form *a-b(a)-rə-j* ‘this very’ (lit.: DEF-this-EMPH-SG) → the presentative form *a-ba-r* ‘here is’ (lit.: DEF-EMPH-this), as in *abar kəlxəz* ‘here is a kolkhoz’ (Genko 1955:102-104). Historically, all these *b* morphemes may go back to a kind of locative verb meaning something like ‘be (somewhere)’.

- (37) a. *jə-w-fa-j-a?*
 PTCP.ABS-2SG.M[A]-eat[AOR]-3SG.NON<HUM>.S-IPW
 ‘What did you eat? (lit.: ‘The one that you ate, it is who/what?’)’
 (Genko 1955:106)
- b. *jə-w-fa-d-a?*
 PTCP.ABS-2SG.M[A]-eat[AOR]-3SG.HUM.S-IPW
 ‘Who did you eat? (lit.: ‘The one that you ate, s/he is who/what?’)’
 (Genko 1955:106)
- (38) a. *jə-ca-wa-j-a?*
 PTCP.ABS-go-PTCP.IPFV-3SG.NON<HUM>.S-IPW
 ‘What is going (there)? (lit.: ‘The one that is going, it is who/what?’)’
 (Genko 1955:107)
- b. *jə-ca-wa-d-a?*
 PTCP.ABS-go-PTCP.IPFV-3SG.HUM.S-IPW
 ‘Who is going (there)? (lit.: ‘The one that is going, s/he is who/what?’)’
 (Genko 1955:107)

As is suggested by the literal translations of (36-38), functionally the interrogative root *-a* in these examples can be perceived as a nominal predicate, which explains the presence of the third person singular intransitive subject markers, *j(ə)-* and *d(ə)-*, and the use of the participle forms *wə-z-zə-pšʃa-wa-* ‘the one that you are looking for’, *jə-w-fa-* ‘the one that you ate’ and *jə-ca-wa-* ‘the one that is going’ instead of the regular finite verbs. This predicative interpretation of *-a* also squares well with the fact that as such it is indifferent to the opposition human vs. non-human.

Besides the bound interrogative root, Tapanta Abaza also has a strategy to form words, functionally equivalent to the interrogative pronominals of languages like English. This strategy is used when no verb is present. It also seems to be recurred to when the question needs to be made more emphatic, somewhat similarly to the possible use of cleft in English in *What is it that you ate?* vs. *What did you eat?* The strategy at issue in Tapanta Abaza consists in the use of some more or less lexicalized non-finite verbal word form, such as *dazəwda* (or contracted *dəzda*) ‘who (exactly)? (SG)’ *dzačʷəja* ‘who? (SG)’, *(j)začʷəja* (rarely *ačʷəja*) ‘what? (SG)’, *začʷəkʷaja* ‘who?, what? (PL)’, *dzapšəja* ‘like who?, resembling who? (SG)’, *(jə)zapšəja* ‘like what?, resembling what? (SG)’, *jəzapškʷaja* ‘like/resembling who?, like/resembling what? (PL)’ (Genko 1955:105-106; Lomtadze 1967b:134).

According to Genko (1955:106), the latter three forms, *dzapšəja*, *(jə)zapšəja*, and *jəzapškʷaja*, are the most transparent for the speakers of Abaza. They all contain the element *apš* ‘(be) like, how’, as in (39) where it is marked with an adverbial participle marker *-ta*. The element *apš* may be further analyzable as *a-*

p-š DEF-EMPH-like, which would be structurally similar to the emphatic form of demonstratives (cf. the footnote 31 in the present section).³²

Tapanta Abaza

(39) *d-šə-nχə-w-z*

3SG.HUM.S-PTCP.ADV.MANNER-work-PTCP.IPFV-PTCP.PST

apš-ta

like[AOR]-PTCP.ADV

‘The way s/he worked...’ (Genko 1955:150)

The final segment *-ja* in *dzapšəja* and *(jə)zapšəja* has the same structure *-j-a* 3SG.NON<HUM>.S-IPW as *-j-a* in (36a, 37a, 38a). In the plural form *jəzapšək^waja*, the *-j-* of *-j-a* may either be 3SG.NON<HUM>.S or 3PL.S, which are identical. The *-z-* preceding *-apš-* is the indirect object or agent participle marker, like *-z-* in (36a) and (36b). The initial *d-* and *jə-* in the first two forms are the third person singular intransitive subject markers, the human *d(ə)-* and the non-human *j(ə)-* respectively, and the *jə* in *jəzapšək^waja* is the third person plural intransitive subject markers, which is indifferent to the human vs. non-human distinction. The element *-k^wa-* in *jəzapšək^waja* is a nominal plural marker. The resulting glosses and literal translations of the three interrogatives *dzapšəja*, *(jə)zapšəja*, and *jəzapšək^waja* are provided in (40a), (40b) and (40c) respectively.

Tapanta Abaza (Genko 1955:106)

(40) a. *d-z-apšə-j-a?*

3SG.HUM.S-PTCP.OBL-like[AOR]-3SG.NON<HUM>.S-IPW

‘like who?, resembling who? (SG) (lit.: ‘The one that s/he is like, it is who/what?’)’

b. *(jə)-z-apšə-j-a?*

3SG.NON<HUM>.S-PTCP.OBL-like[AOR]-3SG.NON<HUM>.S-IPW

‘like what?, resembling what? (SG) (lit.: ‘The one that it is like, it is who/what?’)’

c. *jə-z-apš-k^wa-j-a?*

3PL.S-PTCP.OBL-like[AOR]-PL-3SG.NON<HUM>.S~3PL.S-IPW

‘like who/what?, resembling who/what? (PL) (lit.: ‘The ones that they are like, it is who/what?’ or ‘The ones that they (persons or things) are like, they are who/what?’)’

³² The initial *a-* would then be a definite marker, used to form all demonstratives and identical to the 3SG.NON<HUM> possessor/agent/indirect object affix *a*. The *-p-* would be the emphatic marker or copula *-b-* devoiced to the following *-š-* ‘like, how’. The latter morpheme is also found as the adverbial participle marker *-š(ə)-* in (39) or as the *š-* in *š-pa* ‘how?’ vs. *an-ba* ‘when?’.

The complex structure of these interrogatives is further supported by the fact that according to Genko (1955:106), they may incorporate any noun, as illustrated in (41) with *kət* ‘village’.

Tapanta Abaza (Genko 1955:106)

- (41) *a-rə-j* *a-kət*
 DEF-this-SG DEF-village
jə-z-apš-kət-j-a?
 3SG.NON<HUM>.S-PTCP.OBL-like[AOR]-**village**-3SG.NON<HUM>.S-IPW
 ‘What kind of village is this village?’

The interrogatives *dzačʷəja* ‘who? (SG)’, *(j)začʷəja* ‘what? (SG)’, and *začʷəkʷaja* ‘who?, what? (PL)’ are structurally similar to the three interrogatives just discussed. They differ only in that they use the element *ačʷ(ə)* instead of *apš*. The element *ačʷ(ə)* is constructed of *a-*, the definite marker or the 3SG.NON<HUM> possessor/agent/indirect object affix, and *čʷ(ə)*, which elsewhere functions as a possessive pronominal formative ‘belonging to’, as in *l-čʷə* ‘hers’ (3SG.F.HUM.OBL-belonging.to) (Lomtadze 1967b:134), a destiative nominalizer, as in *fə-čʷə* ‘something to be eaten, food’ (Genko 1955:115), or a verbal root ‘belong to’, as in (42).

Tapanta Abaza (Genko 1955:105)

- (42) a. *də-z-čʷə-w-d-a?*
 3SG.HUM.S-PTCP.OBL-belong.to-PTCP.IPFV-3SG.HUM.S-IPW
 ‘Who does s/he belong to? (lit.: ‘The one that s/he belongs to, s/he is who/what?’)’
 b. *jə-z-čʷə-w-d-a?*
 3SG.NON<HUM>.S-PTCP.OBL-belong.to-PTCP.IPFV-3SG.HUM.S-IPW
 ‘Who does it belong to? (lit.: ‘The one that it belongs to, s/he is who/what?’)’

Thus, *a-čʷ(ə)* can be literally translated as something like ‘its (part, possession, “belonging”)’ (parallel to *l-čʷə* ‘hers’). Consequently, the interrogatives *dzačʷəja* ‘who? (SG)’, *(j)začʷəja* ‘what? (SG)’, *začʷəkʷaja* ‘who?, what? (PL)’ literally mean something like ‘the one to which she/he/it/they relate as its, it is who/what?’, or as Genko (1955:105) paraphrases it ‘to what kind does it/she/he (do they) belong?’.³³

The complex structure of these interrogatives is further supported by the fact that according to Genko (1955:106), they may incorporate any noun, as

³³ The difference between the literal meanings of *d-z-a-čʷə* in *dzačʷəja* ‘who? (SG)’ and *də-z-čʷə-w-* in *dəzčʷəwda* in (42a) is somewhat difficult to render.

illustrated in (43) with *kət* ‘village’, which may be compared to (41) above.

Tapanta Abaza (Genko 1955:106)

- (43) *a-rə-j* *a-kət*
 DEF-this-SG DEF-village
jə-z-a-č^w-kət-j-a?
 3SG.NON<HUM>.S-PTCP.OBL-3SG.NON<HUM>.OBL-belonging.to[AOR]-
village-3SG.NON<HUM>.S-IPW
 ‘What village is this village?’

The lack of differentiation between a human and a non-human meaning in the plural forms *zač^w(ə)k^waja* and *jəzapš^waja* is due to two reasons. First, both in the human and the non-human meaning of these interrogative constructions, the final segment is *-j-a*, literally ‘it is what/who?’, where the third person non-human singular intransitive subject marker *j-* appears to function similarly to the English *it* in *Who is it?* or *What is it?*, which equally does not correlate in (non-)humanness with the interrogative pronominals *who?* and *what?*. This use of *-j-a* in all the interrogative structures at issue inevitably transfers the responsibility for the differentiation between a human and a non-human meaning of the whole construction to the initial third person intransitive subject marker. However, only in the singular does this marker have a separate human form *d(ə)-* and a separate non-human *j(ə)-* (sometimes *Ø-*). In the plural, only one form *j(ə)-* (sometimes *Ø-*), indifferent to the distinction human vs. non-human, is available. Second, the nominal plural marker *-k^wa-* is also indifferent to the human vs. non-human distinction (cf. Genko 1955:124; Chkadua & Klychev 2001:135)³⁴ and thus cannot help to differentiate the two meanings.

By way of conclusion, it may be worth mentioning that the other major dialect of Abaza, Ashkaraua Abaza, avoids the lack of differentiation between ‘who?’ and ‘what?’ through a consistent use of the final *-j-(a)* in similar constructions interrogating about non-humans and the final *-d-(a)* when interrogating about humans. Thus, Ashkaraua Abaza has *z-ak^wə-d-a* ‘who?’ and *z-ak^wə-j* ‘what?’, where *ak^w(ə)* is the verb ‘exist, be’ (Genko 1955:106).

³⁴ Note in this respect that Chkadua & Klychev (2001:135) suggest that originally *-k^wa-* was a non-human plural nominal marker (opposed to the human plural nominal marker *-č^wa-*), which, however, “in contemporary Abaza is getting more and more widely used with human nouns as well”. According to Genko (1955:115-117, 124), *-k^wa-* is indifferent to the opposition human vs. non-human, while *-č^wa-* is indeed restricted to humans, but it is not productive anymore. His description was finished in 1934 (see the editorial note in Genko 1955:3), which suggests that either the use of *-k^wa-* both with human and non-human nominals is older than Chkadua & Klychev (2001) think, or that the two descriptions deal with two different subdialects of Tapanta Abaza.

3.2.2.2 (Standard) Abkhaz

Abkhaz is spoken by more than one hundred thousand people, primarily in the Abkhazia Republic of Georgia. It has three major dialects, Abzhywa (or Abzhui), Bzyb and Sadz. The standard language has been formed on the basis of the Abzhywa dialect.

Standard Abkhaz is similar to Abaza in distinguishing between the bound interrogative morphemes, human *-d-a* (44a) and non-human *-j < *-j-a* (44b), and the verbal constructions functioning as independent interrogative pronominals.

Abkhaz (Klychev & Chkadua 2001:127)

(44) a. *j-ca-d-a?*

PTCP.ABS-go[AOR]-3SG.HUM.S-IPW

‘Who went? (lit.: ‘The one that went, s/he is who/what?’)’

b. *j-ce-j?*

PTCP.ABS-go[AOR]-3SG.NON<HUM>.S-IPW

‘What went? (lit.: ‘The one that went, it is who/what?’)’

Like in Abaza, in standard Abkhaz the differentiation between the human and non-human meanings of the verbal constructions functioning as independent interrogative pronominals largely relies on whether the intransitive subject marker used is human *d(ə)-* or non-human *j(ə)-* (cf. Genko 1955:106; Lomtadze 1967a:112). Consider, for instance, the following forms: *də-z-wə-s-d-a* or *də-z-u-s-t(a)-d-a* ‘who?’ (*wə* is the verb ‘do, act’),³⁵ *jə-z-wə-s-k^{hw}a-d-a* ‘who? (PL)’ (Genko 1955:106; Lomtadze 1967a:112). Here, like in Ashkaraua Abaza, the lack of differentiation between ‘who?’ and ‘what?’ is avoided through a consistent use of the final *-d(a)* when interrogating about humans. However, Abkhaz also has *(jə-)z-ak^wə-(za)-j* ~ *j-z-ak^wu-(ze)-j* ‘what?’ (Genko 1955:106; Klychev & Chkadua 2001:123) and the corresponding human form *d-z-ak^wu-(ze)-j* (< *d-z-ak^wə-(za)-j*) ‘who?’ (Lomtadze 1967a:112),³⁶ which are very much like the Tapanta Abaza forms *dzač^wəja* ‘who? (SG)’ and *(j)zač^wəja* ‘what? (SG)’ discussed above in that they have the final segment *-j-a*, literally ‘it is what/who?’, both in the human and the non-human meanings. Also like Tapanta Abaza, standard Abkhaz uses a single third person plural intransitive subject marker *j(ə)-* indifferent to the opposition human vs. non-human (cf. Lomtadze 1967a:113). Furthermore, like the Tapanta Abaza plural suffix *-k^wa*, the Abkhaz

³⁵ Compare Tapanta Abaza *d-a-zə-w-d-a* ‘who?’.

³⁶ Apparently, the element *ak^w(ə)* ~ *ak^wu* in the Abkhaz forms is cognate to the Abaza verb *ak^w(ə)* ‘exist, be’, as in the Ashkaraua Abaza forms *z-ak^wə-d-a* ‘who?’ and *z-ak^wə-j* ‘what?’ mentioned above. It should be mentioned, however, that the regular form of the verb ‘be’ in Abkhaz is *qa* (cf. Klychev & Chkadua 2001:126; Lomtadze 1967a:115).

plural suffix *-k^{hw}a* is also indifferent to the opposition human vs. non-human, at least when used with verbal roots (cf. Klychev & Chkadua 2001:117).³⁷ This may suggest that just like Tapanta Abaza, standard Abkhaz may also have an undifferentiated ‘who?, what?’ wordform in the plural, something like (*jə*-)z-*a^wkə-k^{hw}a-(za)-j* or (*jə*-)z-*a^wkə-k^{hw}a-(ze)-j*. Unfortunately, neither Lomtadze (1967a) nor Klychev & Chkadua (2001) provide plural forms of the relevant interrogative constructions. Lomtadze (1967a:112) only adds “etc.” after the singular forms, which suggests that the plural forms are possible. The existence of the plural form *z-a^wkə-k^{hw}a-ze-j* (< **j-z-a^wkə-k^{hw}a-ze-j*) was confirmed by Larisa Boyd (p.c.). However, according to her, this form is used only in questions about plural non-humans, while in questions about plural humans one would rather use *z-u-s-t(a)-d-a* ‘who?’ (< **j-z-u-s-t(a)-d-a*), which is the plural of the interrogative *də-z-u-s-t(a)-d-a* ‘who?’ (*wə* is the verb ‘do, act’), already mentioned above.

Besides the forms just discussed, standard Abkhaz also uses an interrogative pronominal root (or a stative verb) *-arban* ‘(be) what?, (be) who?’,³⁸ which obligatorily takes an intransitive subject prefix, e.g. *d-arban* ‘who is s/he?’, *j-arban* ‘what is it?’, *s-arban* ‘who am I?’, *w-arban* ‘who are you? (M.SG)’ (Genko 1955:106; Klychev & Chkadua 2001:123; Lomtadze 1967a:112). The plural form *j-arban* ‘who/what are they? (PL)’ indifferent to the opposition human vs. non-human would normally be expected. The existence of the plural form *j-arban*, homonymous to the singular non-human interrogative *j-arban* ‘what is it?’, was confirmed by Larisa Boyd (p.c.).³⁹ Unlike with the plural *z-a^wkə-k^{hw}a-ze-j* mentioned above, the plural interrogative *j-arban* may be used both in questions about plural non-humans and plural humans. Thus, the interrogative *j-arban* answered both by a singular or plural non-human noun, e.g. *a^wla* ‘a tree’ or *a^wla-k^{hw}a* ‘trees’, and a plural human noun, e.g. *awaa* ‘people’.

3.3 Dravidian languages

The Dravidian linguistic family consists of four branches: Southern, South-Central, Central and Northern. As far as I can judge from the available data, languages with interrogative pronominals used indiscriminately as ‘who?’ or ‘what?’ can be found in every branch, although most typically among the South-Central and Central Dravidian languages. However, it is not always easy to say

³⁷ On nouns, the plural suffix *-k^{hw}a* in Abkhaz is mostly used with non-humans, whereas with humans the plural suffix *-c^{hw}a* is preferred (cf. Klychev & Chkadua 2001:117).

³⁸ The etymology of *-arban* is not clear.

³⁹ Here again, both Genko (1955:106) and Lomtadze (1967a:112) just put “etc.” after the singular forms.

with certainty whether a given Dravidian language really has ‘who?, what?’ interrogative pronominal(s). The reason is that the systems of interrogative pronominals of Dravidian languages are organized in a rather complex way, but most sources I consulted treat them somewhat too cursorily, confining themselves to lists of forms with only a minimum on examples or comments on the patterns of use. The complexity of the Dravidian interrogative pronominal systems is largely due to the complexity of the gender systems of these languages. Therefore, it will be useful to begin with a discussion of the Dravidian gender systems in Section III.3.3.1 and then proceed to a general presentation of the Dravidian interrogative pronominal systems in Section III.3.3.2. In Section III.3.3.3, I will discuss the languages that may prove to have ‘who?, what?’ interrogative pronominals. In all cases, this indiscriminate use of an interrogative pronominal is due to the conflation of certain gender-number distinctions: usually the neuter with the feminine, rarely also with the masculine (in the singular and/or in the plural). Sometimes, the lack of differentiation between ‘who?’ and ‘what?’ seems to be confined to predicative use. By way of conclusion, in Section III.3.3.4 I will briefly present Andronov’s (1978:275-279) reconstruction of the Proto-Dravidian interrogative pronominal system and I will try to elaborate it on certain points, such as the semantics of the reconstructed etymons.

3.3.1 *Dravidian gender systems*

Gender is found in most Dravidian languages. Gender assignment is largely semantic. With only a minor number of exceptions in every language, it is based on the opposition human vs. non-human and for humans on sex, i.e. male vs. female.⁴⁰ This gives us three basic genders, conveniently called *masculine*, *feminine* and *neuter*. However, this ideal threefold distinction is very rarely found as such. Dravidian languages tend to neutralize either the distinction between the masculine and the feminine agreement patterns in one non-neuter agreement pattern (also known as “epicene”, Andronov 1978:170) or between the feminine and the neuter agreement patterns in one non-masculine agreement pattern. Some languages have lost gender distinctions altogether. Gender also often interacts with number, so that in the singular more agreement patterns may be available than in the plural (the opposite is considerably less frequent), or when the number of agreement patterns is the same, it may happen that agreement patterns have coalesced differently in the two numbers (e.g., masculine vs. non-masculine










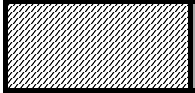




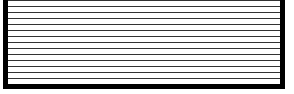

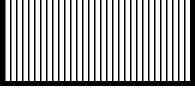




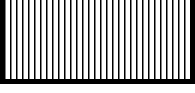
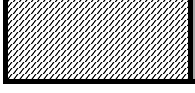


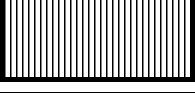

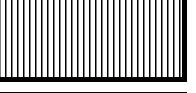
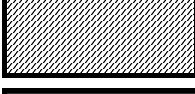

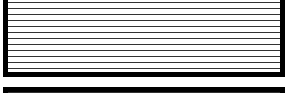
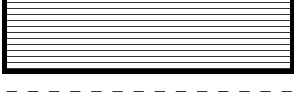
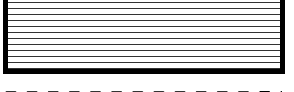


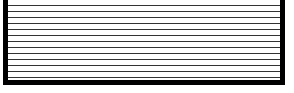




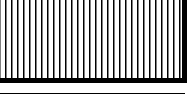
⁴⁰ Sometimes, the terms *rational* and *non-rational* are used instead of *human* and *non-human* respectively, because the names of deities may sometimes be treated like human nouns, while the words like ‘child’ may be treated like non-human nouns.

in the singular, but non-neuter vs. neuter in the plural).⁴¹ In some languages the situation is further complicated by the fact that various targets may differ in the number of genders they distinguish. For instance, numerals, adjectives and demonstrative-cum-anaphoric pronominals may make fewer distinctions than verbs.⁴² According to Andronov (1978:169-173), all in all seven types of gender systems can be found in Dravidian language, plus one type comprising languages that lack grammatical gender altogether. These types are presented in Table 11. However, at least two more types seem to be possible. Types IX and X may need to be added to those mentioned by Andronov. It should also be kept in mind that Andronov's classification is based on the agreement patterns distinguished on verbs. This is why languages like Kuvi, Malayalam or Malpaharia Malto, which make fewer distinctions on verbs than, for instance, on demonstrative-cum-anaphoric pronominals (see note ^{xiv}) to the table in Appendix E and footnote 42 in this section), are classified as types VI, VIII and IX and not as IV, II and III respectively. Similarly, the gender-number distinctions made on verbs do not always coincide with the distinctions made on interrogative pronominals.

⁴¹ In some languages, in the singular the masculine and the feminine agreement patterns may co-exist with a non-neuter one, but the latter is typically honorific at the same time (e.g., in Modern Tamil). Therefore, it would probably be more correct to say that honorificity blocks the expression of gender, just as number sometimes does, rather than to speak of a separate non-neuter gender. Furthermore, it seems that in this case there are no nominal controllers that are lexically specified as (singular) non-neuter.

⁴² The opposite occurs less frequently. An extreme case can be found for instance in Malayalam, where demonstrative-cum-anaphoric pronominals distinguish M.SG, F.SG, N.PL and NON(N).PL, as well as several honorific forms, whereas no gender agreement is marked on verbs (Asher & Kumari 1997; Pillai & Kothandaraman 1972; Andronov 1978:170-171). Another interesting case is represented by Malpaharia Malto (Steever 1998b). In this language the feminine is distinguished from the neuter only in singular demonstrative-cum-anaphoric pronominals (in the accusative, instrumental, ablative and locative cases). A possible solution would be to consider the pronominals both in Malayalam and Malpaharia Malto not as targets but as controllers. This way the distinction between various "gender" forms of demonstrative-cum-anaphoric pronominals would be transferred to the lexical level (e.g., in Malpaharia Malto the distinction would be between a word meaning 'this woman, she' and a word meaning 'this thing, it'). Consequently, the opposition between the non-masculine and masculine genders in Malpaharia Malto or lack of any gender in Malayalam would remain unchallenged. Note that the situation in Malayalam is in fact very similar to that in English. As Corbett (1991:170) points out "whether or not gender is recognized in languages like English [...] depends on [...] one's view of agreement".

Table 11. Dravidian gender systems, agreement on the verb (adapted from Andronov 1978:169-173, 359; additional information comes from: Das 1973; the papers in the volume edited by Steever 1998c and Christiane Pilot-Raichoor, p.c.)

Type	Languages	SG			PL		
		M	F	N	M	F	N
I	Tamil “lower castes” dialects						
II	Tamil (most dialects), Kota, Kannada (most dialects), Badaga, Kodagu, Tulu						
III	Mapali Malayalam, Havyaka Kannada						
IV	Pengo, Kurux (“woman to woman” speech)						
V	Koraga, Bellari, Telugu, Kurux, (Kumarbhag?) Malto						
VI	Kolami, Naiki, Parji, Gadaba, Gondi, Konda, Kui, Kuvi, Sauria Malto						
VII	A Southern Kanara Tamil dialect						
VIII	Malayalam (most dialects), Toda, Yerukala (Kurru), Kuru(m)ba, Brahui						
<hr/>							
?IX	Malpaharia Malto						
?X	Manda						

3.3.2 Dravidian interrogative pronominal systems

Interrogative pronominals from various Dravidian languages are summarized in Appendix E. Typically, Dravidian languages have two sets of interrogative pronominals: the interrogative pronominals distinguishing number-gender, which I call *Set I*, and the interrogative pronominals distinguishing (only) humanness, which I call *Set II*.⁴³ Set I and Set II are often referred to in the literature as

⁴³ I put *only* in brackets here for the following reasons. First, in some languages, number

“selective” and “general” interrogatives respectively (cf. Asher & Kumari 1997:267), with selective implying the gloss ‘which one(s)? (man, woman, thing, etc.)’ and general implying ‘who?’ or ‘what?’. However, such a description of the two sets oversimplifies the real patterns of their use. Even if this description may indeed be largely applicable to certain languages, particularly Southern Dravidian ones, it is clearly misleading when applied to the family as a whole. Consider, for instance, examples (45-48) from South-Central and Northern Dravidian languages in which Set I interrogatives appear to be used as non-selective interrogative pronominals.

Abujhmara Gondi (South-Central Dravidian; Natarajan 1985:128)

- (45) *iw* *bo:rlɔ:r-a:ŋ* *lo:hk*
 this.PL.NON<M>.NOM IPW[S1].PL.M-GEN house.PL.NON<M>.NOM
a:nd-u:ŋ?
 be.PRS-3PL.NON<M>
 ‘Whose (PL) houses are these?’

- (46) *ad* *bad-e:n* *tit?*
 that.SG.NON<M>.NOM IPW[S1].SG.NON<M>-ACC eat.PST.3SG.NON<M>
 ‘What did she eat?’

Malpaharia Malto (Northern Dravidian; Steever 1998b:368)

- (47) *ni:m* *ike-r?*
 2PL.NOM IPW[S1].NOM-2PL
 ‘Who are you(PL)?’

distinction may be marginally available for these interrogatives, if not always in the forms of the interrogatives themselves, then at least in the agreement markers on the verbs (cf. the table in Appendix E and notes thereby). Second, given that Set II interrogatives require a particular agreement pattern on such targets as, for instance, verbs, strictly speaking they do belong to a certain gender. What is important, however, is that for Set II interrogatives humanness primes over gender/number (for Set I interrogatives the situation is the other way around). This is particularly evident in languages like Kui and Kurux. Thus, the Set II human interrogatives in these two languages, *imbai/embai/ba:i* in Kui and *ne:* in Kurux, are followed by a verb with a NON<M>.SG marker (Winfield 1928:45; Grierson 1906:IV); their non-human counterparts trigger the same agreement pattern. The agreement idiosyncrasies of Set II interrogatives are primarily due to the fact that Set II interrogatives are Set I interrogatives in origin. For instance, the Kui human Set II interrogative *imbai/embai/ba:i* ‘who?’ can be compared to the Set I NON<M>.SG interrogative *imbai* in the closely related language Kuvi (cf. the table in Appendix E).

Kuvi (South-Central Dravidian; Reddy 1979:143-144)

- (48) *ni taggu maṇisi emminasi?*
 2SG.GEN dispute man.SG.M.NOM IPW[S1].SG.M.NOM
 ‘[The orphan boy said “I will bring a mediator”, and wandered in search of an intercessor ... The orphan went back to the king’s mediators. They asked the boy:] Who is your mediator? [The boy said “Please wait for a moment, he is coming.”]’

In these examples, the interrogatives belong to Set I but can hardly be characterized as selective. Actually, even in Southern Dravidian languages Set I interrogatives cannot always be characterized as selective, as illustrated in (49) and (50).

Tamil (Southern Dravidian; Pillai & Kothandaraman 1972:351)

- (49) *inta-p puttakam e:tu?*
 this.ADJ-ACC book.SG.N.NOM IPW[S1].SG.N.NOM
 ‘Where/how did you get this book from?’ or ‘Whose book is this?’
 (perhaps, an English translation better covering this range of meanings would be something like ‘What book is this?’, even though from a syntactic point of view it is not exactly parallel to the Tamil original)
- (50) a. *evan va-nt-a:n?*
 IPW[S1].SG.M.NOM come-PST-3SG.M
 ‘Who came? (The speaker knows that the person came is a male)’
 b. *eval va-nt-a:!??*
 IPW[S1].SG.F.NOM come-PST-3SG.F
 ‘Who came? (The speaker knows that the person came is a female)’
 c. *evan va-nt-a:r?*
 IPW[S1].SG.HON(NON<N>).NOM come-PST-3SG.HON(NON<N>)
 ‘Who came? (The speaker knows that the person came is a respectable person)’

In (50), for instance, the use of Set I interrogatives just shows that “the speaker knows that the person referred to is male [(50a)], [...] female [(50b)] [or] respectable [(50c)]” (Pillai & Kothandaraman 1972:351). Admittedly, Set I interrogatives may be the preferred ones in selective contexts, but this is so just because in such contexts the speaker normally knows the number/gender of the members of the set from which a selection needs to be made, whereas in other, non-selective kinds of contexts this knowledge is frequently lacking.

The characterization of Set I interrogatives as selective becomes particularly problematic in South-Central Dravidian languages, because in various languages of this branch Set I interrogative pronominals are further subdivided into the so-called “definite” and “indefinite” interrogative pronominals, with “definite”

actually standing for selective. However, the terms “definite” and “indefinite” are rather misleading as well, because as Bhat (1989:346) points out, “(in)definiteness” does not have to do with referents here but with what he calls “domain of interrogation”. The so-called “definite” interrogative pronominals are usually glossed with the help of an attributive ‘which?’ (e.g., ‘which man, woman, etc.?’), which implies that they should be selective. The so-called “indefinite” interrogative pronominals are usually glossed with an attributive ‘what?’ (e.g., ‘what man, woman, thing, etc.?’) and sometimes with ‘what kind of [N]?’ (e.g., ‘what kind of man, woman, thing, etc.?’),⁴⁴ which implies that they should be non-selective.

That Set II interrogatives are not very different from Set I interrogatives becomes particularly clear when Set II interrogatives are examined from a diachronic point of view. As can be readily observed from comparison of the forms of the two sets in Appendix E, Set II interrogatives seem to be always Set I interrogatives in origin. Usually, they even preserve their original agreement pattern, which from a synchronic point of view may often look rather idiosyncratic, as for instance the NON<M>.SG agreement pattern of the human Set II interrogatives ‘who?’ in Kui and Kurux (cf. footnote 43 in this section and note ^{xx}) to the table in Appendix E). In other words, Set II interrogatives are just Set I interrogatives that at a certain point happened to be chosen as default human or non-human interrogatives. Thus, Set II human interrogatives meaning ‘who?’ usually go back to a Set I PL.NON<N> interrogative (e.g., in Tamil), a Set I SG.M interrogative (e.g., in Abujhmari Gondi), or Set I PL.M interrogative (e.g., in Muria Gondi). Set II non-human interrogatives meaning ‘what?’ may go back to a Set I SG.N or SG.NON<M> interrogative (e.g., in Pengo, Naiki and Modern Telugu). Interestingly, in the case of ‘what?’ the source is often a predicative Set I SG.N or SG.NON<M> interrogative, or its attributive derivative meaning ‘what (kind of) [N]?’ For instance, the Naiki Set II non-human interrogative *ta:n* ‘what?’ may be compared to the predicative form *ta:n-ed/tan-d* of the Set I SG.NON<M> interrogative *e(:)d*.⁴⁵ Somewhat similarly, the Modern Telugu Set II

⁴⁴ As far as I can judge, at least in Kui the ‘what kind of?’ gloss must be rare, since this meaning is expressed with the interrogative *isingi* ‘how?’ used as a modifier, as in *isingi tari* ‘what kind of thing/woman?’ (Winfield 1928:51). I do not have enough data for the other South-Central Dravidian languages. Note that the use of ‘how?’ as an adnominal modifier with the meaning ‘what kind of [N]?’ is not confined to South-Central Dravidian languages. For instance, Christiane Pilot-Raichoor (p.c.) reports the same kind of use for the interrogative *etta-* ‘how?’ in the Southern Dravidian language Badaga.

⁴⁵ Most likely, the first part *ta(:)n* of the predicative form goes back to the widespread Dravidian third person singular “reflexive” pronoun of the same form. According to Steever (1998d:22), even though the reflexes of **ta:n* (SG)/ **ta:m* (PL) function as reflexive pronouns in

non-human interrogative *e:mi* ‘what?’ goes back to an Old Telugu Set I SG.N special predicative interrogative *e:mi* (see note ^{vii}) to the table in Appendix E for more details). In all probability, various Southern Dravidian Set II non-human interrogatives of the forms *entV/en(nV)/ettV* and the like are attributive in origin. In fact, many of these interrogatives still function attributively as well, the meaning being ‘what (kind of) [N]?, [N] like what?’. Their original form is probably something like **en-ta*, or less likely **em-ta* (the reconstruction with the labial nasal is due to Andronov 1978:278), sometimes further extended with a singular neuter gender marker **-ttu*. Here, the first part may be compared to the Set I “indefinite” interrogative root *Vn-* of various South-Central Dravidian languages (as for instance *an-/in-* of Kui) or the Telugu interrogative *e:mi*. As to the second part *-ta* (or the like), it shows up as an adjectivizing suffix in various Dravidian languages. For instance, in Modern Tamil attributive demonstratives have a similar structure *Vn-ta/Vn-na* (Pillai & Kothandaraman 1972:353-354, Andronov 1978:279) and in Abujhmara Gondi *-ta:* is a productive adjectivizing suffix, as in *bane:k* ‘beauty’ → *bane:k-ta:* ‘beautiful’ (Natarajan 1985:147-149).

It seems that typically the differentiation between the interrogatives of Set I and Set II is organized in a privative way. That is, there are contexts (semantic or morphosyntactic) where the interrogatives of both sets are in principle possible and there are contexts (semantic or morphosyntactic) where the interrogatives of only one of the two sets are possible. In the former case, Set II interrogatives can be seen as more “general” indeed. Set II interrogatives only indicate the speaker’s assumptions on whether the element in the scope of interrogation is human or non-human, whereas Set I interrogatives specify the assumptions on its number-gender. Thus, the Set I interrogatives in the Tamil examples (50a), (50b) and (50c) can all be replaced, in principle, with one human Set II interrogative *ya:r* ‘who?’, as in (51).

Tamil (Southern Dravidian; Pillai & Kothandaraman 1972:351)

- (51) *ya:r* *va-nt-a:r (-a:rkaḷ, -a:n)?*
 IPW[S2].HUM.NOM come-PST-3SG.HON(NON<N>) (-3PL.NON<N>, -3SG.M)
 ‘Who came? (The speaker knows that somebody came, this somebody may be a male or female person)’⁴⁶

Some contexts where the interrogatives of only one set are possible recur in language after language. For instance, many languages seem to strongly prefer

most modern Dravidian languages, originally these must have been third person anaphoric pronouns.

⁴⁶ Pillai & Kothandaraman (1972:351) themselves use only the 3SG.HON[NON<N>] marker *-a:r* in this example; see note ⁱⁱⁱ) to the table in Appendix E on the possibility of the other two markers, 3PL.NON<N> *-a:rkaḷ* and 3SG.M *-a:n*.

non-human Set II interrogatives to the respective Set I neuter interrogatives in situations when the question is about an action and not an entity (52).

Kannada (Southern Dravidian; Bhat 1989:368)

- (52) *avanu e:nu ma:d-id-a?*
 3SG.M.NOM IPW[S2].NON<HUM>.NOM do-PST-3SG.M
 ‘What did he do?’

Other contexts are language specific. For instance, in Malpaharia Malto (Steever 1998b:368), Set I interrogative *ike-* is not used in oblique cases to inquire about humans. Instead the oblique cases of Set II human interrogative *ne:d(u)* ‘who?’ should be used. At the same time, only *ike-*, not *ne:d(u)*, may be plural and only *ike-*, not *ne:d(u)*, may be used as a predicate nominal.

3.3.3 Dravidian languages with ‘who?, what?’ interrogatives

Several Dravidian languages may lack formal differentiation between the Set I interrogative pronominals questioning about humans (i.e., masculine and feminine) and those questioning about non-humans (i.e., neuter). In all cases, this lack is due to the conflation of certain gender-number distinctions in these interrogative pronominals.

However, it would be wrong to think that when there is a Set I interrogative pronominal form which is either both neuter and feminine or indifferent to gender, we are always dealing with a real case of lack of formal differentiation between ‘who?’ and ‘what?’. First, the situation may be complicated by interaction of number-gender with case. That is, the differentiation may be lacking only in certain cases (cf. Kui or Malpaharia and Sauria Malto). Second, even though a given interrogative may be formally non-masculine or even not distinguish gender at all, this does not necessarily mean that it can be used in questions both about women and non-humans or humans and non-humans. Thus, the interrogative may be marked with a non-masculine gender-number or non-masculine gender-number-case marker, as for instance Sauria Malto *ne:(ri)-θ*, Parji *na:t-ot* or Muria Gondi *b-odd*. However, it seems that in the first case, the interrogative can only be used in questions about women and in the second only about non-humans. The situation in Muria Gondi is not clear. Furthermore, even without being marked with a special non-masculine gender-number(-case) marker, a given interrogative may require a non-masculine agreement pattern on the verb when used as a subject, as for instance Kui *imbai* or Kurux *ne:*. Note, however, that both interrogatives can be used only in questions about humans.

Unfortunately, the sources I consulted usually provide only forms of interrogatives with a very limited number of examples (let alone any information on the larger context) and very rarely examples with the interrogatives not

distinguishing the neuter from the feminine (or the feminine and the masculine).⁴⁷ What is particularly deplorable, the latter kind of interrogatives is hardly ever illustrated with examples where any ambiguity might be possible between the human and non-human meanings (at least when a larger context is not taken into consideration), such as for instance ‘IPW.SG.NON⟨M⟩ is 3.SG.NON⟨M⟩?’, that is ‘Who/what is this (woman/thing)?’, or ‘IPW.SG.NON⟨M⟩ is.lying-3.SG.NON⟨M⟩ there?’, etc., or where the distinction between the human and non-human meanings depends on a different element within the same clause than the interrogative, as for instance in ‘IPW.SG.NON⟨M⟩ is your wife/job?’.

Due to their somewhat cursory treatment, the exact patterns of use of interrogative pronominals often are not fully clear in most sources I consulted either. For instance, it is often difficult to know for sure whether a given interrogative is or is not confined to selective contexts. Moreover, as has been demonstrated in Section III.3.3.2, the claims on the exclusively selective nature of the interrogatives of Set I, often prove to be misleading on closer examination. This is quite important, because if such an interrogative were indeed confined to selective contexts, then it would not represent the kind of lack of formal differentiation that we are looking for, i.e. a lack of formal differentiation between the non-selective interrogative pronominals ‘who?’ and ‘what?’. Clearly, the best way to sort out all such cases would be to go and gather first-hand data from native speakers.

Taking into account the limitations mentioned above, I will not make any strong claims on lack of formal differentiation between ‘who?’ and ‘what?’ in a given language. For the time being, the only reasonable thing to do, I believe, would be to make up a list of languages that may later prove to have ‘who?, what?’ interrogative pronominals. Such a list is presented in Table 12, the forms of possible ‘who?, what?’ interrogatives are given as well. However, judging by the data available, certain languages seem to have higher chances than others of really possessing such a ‘who?, what?’ interrogative. I marked such “best candidates” with an asterisk. My choice of the “best candidates” has been determined by several factors. First, somewhat less equivocal examples may be available. For instance, Andronov (1980:53) defines the Brahui interrogative *ara:(d)* (SG)/*ara:fk* (PL) as “‘who?’, ‘what?’, ‘what kind of?’, ‘which?’ [...] used both as a substantive and as an adjective” and provides (53-55) as examples of its

⁴⁷ Admittedly, texts are sometimes available. However, they are very rarely dialogues, which implies that questions are underrepresented there. Furthermore, glosses or even just word-by-word translation are rarely provided.

Table 12. Dravidian languages which may have a ‘who?, what?’ interrogative pronominal

Language	Possible ‘who?, what?’ interrogative pronominals
Toda*	<i>e:θ(u)</i> (SG), <i>e:θa:m</i> (PL) (in predicative use)
Old Telugu	<i>e:di</i> (SG.NON⟨M⟩)
Modern Telugu	<i>e:di</i> (SG.NON⟨M⟩)
Muria Gondi	<i>bodd, bad</i> (SG.NON⟨M⟩); <i>bovv(es(k))</i> (PL.NON⟨M⟩)
Abujhmaria Gondi*	<i>bad</i> (SG.NON⟨M⟩); <i>baw</i> (PL.NON⟨M⟩)
Konda*	<i>ayed, inika(d)</i> (SG.NON⟨M⟩); <i>ayek, inikeŋ</i> (PL.NON⟨M⟩)
Kuvi*	<i>emminayi, amba(:)yi, imbai</i> (SG.NON⟨M⟩); <i>emmina²i, amba(:)²i, imbai, imbaari</i> (PL.NON⟨M⟩)
Manda	<i>amne</i> (PL.NON⟨M⟩)
Kolami (a)*	<i>e:d</i> (SG.NON⟨M⟩); <i>edav, e:v</i> (PL.NON⟨M⟩)
Kolami (b)	<i>ed</i> (SG.NON⟨M⟩); <i>edav</i> (PL.NON⟨M⟩)
Naiki*	<i>e:d</i> (SG.NON⟨M⟩); <i>e:v</i> (PL.NON⟨M⟩)
Gadaba*	<i>e:di</i> (SG.NON⟨M⟩); <i>e:vi</i> (PL.NON⟨M⟩)
Malpaharia Malto*	<i>ike-ð</i> (SG.NON⟨M⟩.NOM), <i>ike-r</i> (PL.NON⟨M⟩) (in predicative use)
Brahui*	<i>ara:(d)</i> (SG); <i>ara:fk</i> (PL)

substantive use. Admittedly, given that these examples are presented out of any context, the selective interpretation cannot be excluded and we have to rely on Andronov’s translations.

Brahui (Northern Dravidian; Andronov 1980:53)

- (53) *da: Daggi: ara:na: e?*
 this cow.NOM IPW[S1].SG.GEN be.PRS.3SG
 ‘Whose cow is this?’
- (54) *ara:RaT xalkus-ta?*
 IPW[S1].SG.INS strike.PST.2SG-3SG.OBJ
 ‘What did you strike him with?’
- (55) *da: ka:re:me ara:fte:ki kare:nus?*
 this deed.ACC IPW[S1].SG.DAT do.PST.2SG
 ‘Who (PL) did you do it for?’

In his description of Kuvi, Schulze (1911:71) defines the Set I NON⟨M⟩.SG interrogative *imbai* as ‘which woman or thing?’. However, it also seems possible to use it non-selectively in questions about women (56) or things (57).

Kuvi (South-Central Dravidian; India)

- (56) a. *imbai* *wa:te?*
 IPW[S1].NON<M>.SG.NOM come.PST.3SG.NON<M>
 ‘Who came? (“in reference to a female of lower rank”)’ (Schulze 1911:72)
- b. *imbai* *wa:tu?*
 IPW[S1].NON<M>.PL.NOM come.PST.3PL.NON<M>
 ‘Who came? (“in reference to a female of [...] higher rank”)’ (Schulze 1911:72)
- (57) *ninge imbinani ka:vale?*
 2SG.NOM IPW[S1].NON<M>.SG.ACC want.PRS.2SG
 ‘What do you want?’ (Schulze 1911:73)

Similarly, for Toda Emeneau (1984:99, 103) gives the predicative interrogative pronominals *e:θ(u)* (Set I SG; sometimes also PL, as in Emeneau 1971:690)⁴⁸ and *e:θa:m* (Set I PL; in Emeneau 1971:691-692 also as *e:θum*),⁴⁹ which both are claimed to be selective, viz. ‘which one(s)?’, and opposed to non-selective *o:ry* ‘who?’ (Set II HUM; predicatively also as *o:ryu*) and *in* or *i:* ‘what?’ (Set II NON<HUM>; predicatively also as *inu*). However, I have also found examples of the (predicative) non-selective use of the “selective” Set I interrogatives, both as ‘who?’ (58a) and ‘what?’ (59-60). For the sake of comparison, I also provide example (58b), which immediately follows (58a) in the source and is analogous to it but involves the Set II interrogative *o:ry* ‘who?’ instead of the Set I interrogative *e:θ*.

Toda (Southern Dravidian)

- (58) a. [*Ko:xosyk pi:iθik*] *pu tōwad o:dyip o:t i o:t e:θ?*
 new god that.dances man this man IPW[S1]
 [*iθiθik. Ka:sθo:r pu tōwad o:dyip o:t pi:θko:riy iθiθiŋ*]
 ‘[You went to *To:ro:r* mund. You said:] Which (i.e., ‘who?’) is this man, the man who is dancing as a new god? [The man of *Ka:s* said: ‘The man who is dancing as a new god is *Pi:θko:r*.’]’ (Emeneau 1971:690)

⁴⁸ This interrogative is also claimed to have different human and non-human oblique stems, viz. *e:n-* and *e:θ* respectively. However, I have found no example of these oblique stems in Emeneau’s (1971) collection of Toda texts and his (1984) Toda grammar and texts.

⁴⁹ “There are few occurrences in the prose texts of *e:θ*, all are P[redicates] in S[ubject]P[redicate] constructions” (Emeneau 1984:103).

- b. [I:škityk pi:iθik] ko(ŋ) ɲa:sθ xis o:štyip o:ʔ
 ?song-words beauty making that.composes man
 i o:ʔ o:ry?
 this man IPW[S2].HUM
 [idiθik. Ka:sθo:r ko(ŋ) ɲa:sθ xis o:štyip o:ʔ te:fxwiri:y idiθiŋg]
 ‘[You went to Ka:s mund. You said:] Who is this man who composes
 song-words beautifully? [The man of Ka:s said: ‘The man who
 composes song-words beautifully is Te:fxwir.’]’ (Emeneau 1971:690)
- (59) kisiθ nöw i nöw e:θ?
 that.is.made song this song IPW[S1]
 [idiθik. I pu nöw ko:ly nöwi:y idiθiŋg sondari:z]
 ‘[You said:] Which (i.e., ‘what?’) are these songs, the songs that you
 made? [Sondari:z said: ‘This new song is a riddling song.’]’ (Emeneau
 1971:690)
- (60) i paw e:θ [wi:dyit fa:w? iθoθi pa:w θöwfa:w]
 this river IPW[S1]
 ‘Which (i.e., ‘what?’) is this river, [the river that runs? This is Paikara
 river]’ (Emeneau 1971:717)

Second, even when no examples are provided, the sources may contain descriptions which are difficult to interpret otherwise than as statements on the existence of a ‘who?, what?’ interrogative. In most other cases my choice has been determined by structural considerations. Most importantly, when in a given language the Set II human interrogative ‘who?’ is identical to (i) a Set I SG.M (as, e.g., in Muria Gondi or Gadaba) or (ii) a Set I PL.M interrogative (as, e.g., in Abujhmara Gondi, Kuvi or Naiki), the following question needs to be answered. Can the Set I singular and/or plural *non-masculine* form be used for humans, viz. women, or has the Set I masculine form come to be used instead even when it is known that the person in issue is a woman? I would expect that when a formally PL.M interrogative comes to be chosen as a default human interrogative for situations where the gender-number of the element inquired about are unknown, the plural interrogative (as compared to the formally singular one), would be more readily able to spread even further to the contexts where gender-number are known (and even shown on another element in the clause) but do not correspond to the gender-number value of the interrogative itself. Such a spread would confine the respective NON⟨M⟩ forms of the interrogatives to questions about non-humans and would rule a possibility for them to function as a ‘who?, what?’ interrogative pronominal.

My expectation is due to the following reasoning. Given that the number distinction is particularly important with human referents (as opposed to non-human referents) and that the singular is generally expected to be the unmarked

number (cf. Corbett 2000:17, 70), the use of a plural form instead of the corresponding singular one in the situation where the number-gender of the human referent(s) is indeterminate can in itself be considered as a rather marked choice. Therefore, if the marked form has been preferred to the unmarked one to function as the default choice in the context where in principle both forms could have been expected all things, except markedness, being equal, it would also be natural to expect this marked form to be preferred to the unmarked one to function as the default choice in the context where neither of the two could be expected (i.e., where the number-gender value of each contradicts the one already established for the referent elsewhere in the clause).

3.3.4 On the Proto-Dravidian interrogative pronominal system

As mentioned in Section III.3.3.2, typically Set II interrogatives are Set I interrogatives in origin. Therefore, a reconstruction of the Proto-Dravidian interrogative pronominal system equals a reconstruction of the Set I interrogative pronominal system. Thus, for Proto-Dravidian Andronov (1978:275-279) reconstructs one interrogative base with four alternants **ya:v-/*ya:m-/*ya:ŋ-/*ya:l-*,⁵⁰ which can be further split into the actual interrogative element **ya:-* and four deictic elements: **ya:v- < *ya:-av-*, **ya:m- < *ya:-am-*, **ya:ŋ- < *ya:-aŋ-*, **ya:l- < *ya:-al-*. The interrogative *ya:/e:* (and the like) is still found in various modern Dravidian languages as an attributive interrogative meaning ‘which [N]?', ‘what [N]?', as for instance *ya:* in Kannada or *e:* in Badaga. After the deictic element, a number-gender marker is attached. The combination of a deictic root and a number-gender marker is nothing else but a demonstrative-cum-anaphoric pronominal. It is therefore not surprising that the existence of ‘who?, what?’ interrogatives is always due to the conflation of certain gender-number distinctions. Interestingly, the original construction [interrogative pronominal + deictic-cum-anaphoric pronominal] looks very much like a kind of cleft ‘which.one/what.one (is) he/she/it/they/etc. (that...)?’ or ‘who/what (is) he/she/it/they/etc. (that...)?’, with the interrogative pronominal functioning as a predicate nominal.

According to Andronov (1978:261-269), four series of deictic roots can be reconstructed for Proto-Dravidian, which differ only in their initial vowel, that is **a-* for the meaning ‘distal/absent’, **u-* for ‘medial (closer to the interlocutor)’,

⁵⁰ Note that the form **ya:ŋ-* may also be reconstructed as **ya:n-*, that is with a dental nasal *n* instead of the alveolar nasal *ŋ*. As Andronov (1978:139) points out himself, the reconstruction of a phonemic distinction between the two nasals, *n* and *ŋ*, on the Proto-Dravidian level is problematic due to the lack of precision in the existing descriptions of apical nasals for several languages.

e-* for ‘medial (closer to the speaker)’ and **i-* for ‘proximal (to the speaker)’. It is noteworthy that the interrogative base is built with the ‘distal/absent’ deictic series **av-*/am-*/**an-*/**al-*. Unfortunately, Andronov does not say much about the possible functional distinctions the consonants in these forms might have expressed. He only notes that the “alternation *v/m* is widely represented in Dravidian languages” (1978:268), which seems to imply that **av-*/**am-* were just allomorphs with no difference in meaning. As to the other two forms, **an-* and **al-*, the following can be said at least about their use within the interrogative base, **ya:n-* < **ya:an-* and **ya:l-* < **ya:al-*. The interrogative **ya:l-* seems to have had a locative meaning ‘where?’. At least, all examples that Andronov himself adduces in support of its reconstruction have a locative or similar meaning: “Toda, Naiki *e:l* ‘where (to)?’, Gadaba *e:lcoy* ‘where?’, and probably Telugu *e:la(:)* ‘why?’” (1978:278). Interestingly, the Proto-Dravidian locative case markers reconstructed by Andronov (1978:225-228) also typically have the form **-Vl*.

The interrogative **ya:n-* is best considered in comparison to the interrogative **ya:m-*/**ya:v-*. The first possible explanation is that similarly to **ya:l-*, the two interrogatives may reflect some case marking distinctions. In fact, **-m* is reconstructed by Andronov (1978:225-228) as one of the two nominative case markers (the other one being zero), while **-an* and **-in* are claimed to be the typical sources of the accusative and the genitive markers, respectively. Note that in Dravidian languages the case labelled “nominative” is often used in a rather wide range of functions, with subject marking being only one of them (Andronov 1978:227). Furthermore, the forms that Andronov uses as the primary evidence for the reconstruction of the **Vn-* deictics are oblique bases of demonstrative-cum-anaphoric pronominals. Thus, he mentions the Toda “*an-*, *in-* oblique bases of the pronouns *aθ*, *iθ* ‘he, she, it’” (1978:268).

However, the opposition **ya:m-*/**ya:v-* vs. **ya:n-* may also be interpreted without resorting to case. Thus, originally the distinction might have been between a selective interrogative ‘which [N]?’ and a non-selective ‘what [N]?’, similar to the distinction between “definite” and “indefinite” interrogatives found in several South-Central Dravidian languages (cf. Section III.3.3.2). The data of languages like Kui, Kuvi and Pengo seem to suggest that **ya:m-*/**ya:v-* was the selective interrogative and **ya:n-* non-selective.⁵¹ In most other languages this

⁵¹ Admittedly, only in Pengo the original **-m-*/**-v-* vs. **-n-* distinction between the forms of the selective and non-selective interrogatives has been preserved more or less intact, whereas in the Kui and Kuvi forms it has been somewhat obscured by later phonological developments. For instance in Kui, where the selective and non-selective pronominal interrogatives of Set I, *est-* and *an-/in-*, correspond to the two attributive interrogatives, *esti* ‘which [N]?’ and *ani/ini* ‘what [N]?’ respectively (Winfield 1928:45), the phonological development may have been something

opposition has been lost, most frequently in favour of the selective form **ya:m-*/**ya:v-*. Given the usual correlations between referentiality and case, this interpretation does not necessarily have to exclude the case interpretation of the opposition **ya:m-*/**ya:v-* vs. **ya:n-* proposed above.

3.4 Kusunda

The Kusunda language, as described by Watters (2006), used to be spoken by a group of semi-nomadic hunter-gatherers in central and midwestern Nepal. Today, the language is moribund with no more than a handful of speakers left, who have varying degrees of fluency in it. According to Watters, Kusunda is a language isolate, “the sole survivor of an ancient aboriginal population once inhabiting the sub-Himalayan regions before the arrival of Tibeto-Burman and Indo-Aryan speaking peoples” (2006:10).⁵²

According to Watters (2006:35, 48), Kusunda has two interrogative pronominals *nəti* ~ *nədi* ‘who?, what?’ and *nə(t)n* ‘what?, why?’.⁵³ In what follows I will first illustrate the use of these interrogatives and present their etymology proposed by Watters (2006), who hypothesizes that the two interrogatives are based on the same root marked with two different “old classifier morphemes”. After this, I will present an alternative hypothesis, which suggests that the two interrogatives may be nominalizations of an originally predicative interrogative root meaning ‘be which one?’.

The interrogative pronominals *nəti* ~ *nədi* ‘who?, what?’ and *nə(t)n* ‘what?, why?’ are illustrated in (61-62) and (63-64) respectively. It should be mentioned, however, that (61a) is the only example provided by Watters (2006) for *nəti* as

like **ev-ti* > **efti* > **exti* (or rather **eçti*) > **esti* and **Vnti* > **Vnni* > **Vni*. In Kuvi, where the selective and non-selective pronominal interrogatives of Set I are based on the elements *emm-* and *amb-/imb-*, the development may have been something like **em-b...* > **emm-* and **Vn-b...* > **Vmb-* respectively. Andronov (1978:278) suggests that *-b-* in the forms of interrogative pronominals in various South-Central Dravidian languages is epenthetic in origin, but I strongly doubt that this purely phonological explanation is correct.

⁵² The Ethnologue classifies Kusunda as Tibeto-Burman. Attempts have been made to relate Kusunda to various language families and isolates spoken in the Indian subcontinent, as well as much further away (e.g., Burushaski, Nahali, Austro-Asiatic and Caucasian languages). References to the relevant sources can be found in Watters (2006:10).

⁵³ The two forms are zero-marked nominatives. The other two “syntactic” cases, the genitive and the accusative-dative, are usually marked only for animate referents (Watters 2006:50-53); the markers are *-(y)i/- (y)e* and *-da* respectively. Furthermore, the marked genitive seems to be restricted to the expression of possession.

‘what?’. In all other examples it is human, also when used as an indefinite pronominal as in (65).

Kusunda

- (61) a. *nəti ta?*
 IPW DEM.INAN
 ‘What is this?’ (Watters 2006:48)
- b. *nəti na?*
 IPW DEM.AN
 ‘Who is this?’ (Watters 2006:48)
- (62) *gimi təmbə-g-i nəti?*
 money send-3-PST IPW
 ‘Who sent the money?’, ‘Who is the money sender?’ (Watters 2006:117)
- (63) *nu məbə-n-i gipən nətn?*
 2SG hear-2-PST word IPW
 ‘What is the matter/talk you heard?’ (Watters 2006:85)
- (64) *g^hərun nətn n-əm-əo? k^həŋgu qwon*
 hot IPW 2-eat-NEG.IRR cold drink.IMP
 ‘Why won’t you drink [sic] it hot? (Then) drink it cold!’ (Watters 2006:48)
- (65) *tsi ts-ip-n-da nəti aoda dəi-ən*
 1SG 1-sleep-REAL-SBJV INDF door knock-REAL
 ‘While I was sleeping someone knocked at the door’ (Watters 2006:138)

Example (61a) can be compared to a structurally similar example (61b), where *nəti* is said to mean ‘who?’. The difference between (61a) and (61b) is in the demonstratives used, *ta* “inanimate” and *na* “animate” respectively. Note, though, that at least in combination with the interrogative pronominal the distinction seems to be between human and non-human rather than animate and inanimate. Furthermore, Watters points out that “in many contexts [...] the two terms [i.e., *na* and *ta*] seem to be interchangeable” and “it is possible that *na* is the functionally unmarked member of the two” (2006:49). Indeed, Watters’ (2006) description seems to contain only one example with *ta* (*ta kam* ‘this work’, 2006:58) against quite a few with *na*, almost all of which are with inanimate referents. These demonstratives also seem to be part of the third person anaphoric pronoun, *gina* according to Watters’ own data and *git* according to Reinhard & Toba’s (1970) report (cited via Watters 2006:44). According to Watters (2006:49), *gina* may also be used as a distal demonstrative.

The two interrogatives, *nəti* ~ *nədi* and *nə(t)n*, are obviously related. Thus, Watters (2006:47-48) suggests that *-ti* and *-n* may be “old classifier morphemes”,

which he also believes to occur in the numeral ‘one’ *qasti* ~ *qa(:)sn*.⁵⁴ The two forms of the numeral are not reported to have any difference in meaning. In Watters’ description only examples with the form *qasti* are found, both with inanimate reference (e.g., referring to one rupee, 2006:55) and animate reference (e.g., referring to one child, 2006:88). The presence of *t* in the interrogative *nətn* can be explained in two ways. First, we may assume that in both *nəti* ~ *nədi* and *nə(t)n* the base is *nət-* to which *-ti* and *-n* have been added, i.e. **nət-ti* > *nəti* > *nədi* and *nət-n* > *nə-n* respectively. Second, the base may be *nə*, which would imply two layers of the classifiers, **nə-ti-n* > *nə-t-n* ~ *nə-n*. In fact, the same might be true for the numeral ‘one’ as well, **qas-ti-n* > **qas-t-n* > *qa:s-n*.

As such, the classifier hypothesis looks quite appealing. Furthermore, if we assume that the interrogative *nəti* ~ *nədi* is originally exclusively human ‘who?’, as in fact its use in most contexts seems to suggest, then its occasional use instead of the non-human *nə(t)n* ‘what?’ can be compared to the already mentioned spread of the animate demonstrative *na* to the use with inanimate referents instead of the inanimate demonstrative *ta*. The development of the third person anaphoric pronoun *gina* seems to show the same pattern again (cf. above).

However, what I find somewhat strange about the presumed classifiers is that they do not exist elsewhere in the language and that in fact, no other classifiers or traces thereof are mentioned by Watters (2006). At the same time, Kusunda has two morphemes, *-di*, a verbal adjectivizer/nominalizer (68-70), and *-n*, a verbal realis marker and the marker of a dependent verb (also that of a relative clause) (65-67), which might be the sources of *-ti* and *-n* in the interrogative and the numeral as well.

Kusunda

- (66) a. *hyoq-n* *amba*
 hide(*vi*)-ADJZ animal/meat
 ‘hidden/hiding animal (i.e., animal that has hidden itself/ is hiding)’
 (Watters 2006:116)
- b. **hyoq-n* *gimi*
 *hide(*vi*)-ADJZ money
- (67) a. *yo-ən* *amba*
 cook(*vi*)-ADJZ animal/meat
 ‘cooked meat (i.e., meat that is cooked)’ (Watters 2006:115)

⁵⁴ The form *qa:sn* is given in the dictionary (Watters 2006:149) and the form *qasn* in the text (Watters 2006:48).

- b. *hul-ən* *amba*
 cook(vt)-ADJZ animal/meat
 ‘cooked meat (i.e., meat that has been cooked by somebody)’ (Watters 2006:115)
- (68) *bem-di*
 fall(vi)-ADJZ
 ‘fallen (i.e., the one that has fallen)’ (Watters 2006:118)
- (69) *hol-di* *amba*
 cook(vt)-ADJZ animal/meat
 ‘cooked/cooking meat (i.e., meat that is cooked/ is cooking)’ (Watters 2006:115)
- (70) *kila-di*
 steal(vt)-NMLZ
 ‘thief’ (Watters 2006:118)

This hypothesis implies that both the interrogative and the numeral are predicative in origin, **nə(t)*- ‘be which one? (human or non-human)’⁵⁵ and **qa(:)s*- ‘be one (human or non-human)’ respectively, which were used as attributes and/or nominals when marked by an adjectivizing/nominalizing morpheme.⁵⁶ The *-t* in the interrogative may be the so-called “anticausative”/intransitive marker.⁵⁷ In light of this hypothesis, let us first examine the numeral ‘one’ *qasti* ~ *qa(:)sn* and then the interrogative.

⁵⁵ Perhaps, also ‘be where?’, compare in this respect the formation of the interrogatives in “Khoisan” language Western !Xóǀ discussed in Section III.2.2.2. Note that according to Watters (2006:15), the interrogative ‘where?’ *hampe* ~ *kampe* in Kusunda is a loan from Tibeto-Burman language Yamphu.

⁵⁶ The two cases would not be the only instances of lexicalization of a predicative phrase into a nominal lexeme in Kusunda. For instance, Watters (2006:47) suggests that the word *gidzi* ‘name’ may originate from *g-i-dzi* ‘s/he says’ (3-say-3.REAL).

⁵⁷ This “anticausative” suffix is likely to be related to the locative copula *tsi* ‘be (somewhere)’. Note in this respect that according to Watters (2006:102), for some verbs containing the “anticausative” suffix *-t* it is possible to mark the subject agreement twice, after the verb root as well as after the suffix *-t*. There is also another “anticausative” suffix *-q* for which this is not possible. Furthermore, there is a “causative” suffix *-a/-ə*, as in *hər-a* ‘to open (vt)’ and its “anticausative” derivative *hər-a-q* ~ *hər-ə-q* ‘to open on its own (vi)’ (Watters 2006:117, 143). The “causative” suffix originates in the verb *-a/-ə* ‘make, do’ (Watters 2006:60-63, 97). However, I prefer to put the terms *causative* and *anticausative* used by Watters (2006) in inverted commas because in pairs such as *p^hir-a* ‘to fill (vt)’ vs. *p^hiru-t* ‘to fill (vi)’ (Watters 2006:101, 149), it is strictly speaking impossible to say which verb is derived from which one. Rather, the two are based on the same root and the suffixes just mark the verb as transitive or

Interestingly, next to the numeral ‘one’ *qasti* ~ *qa(:)sn* Kusunda also has an intransitive verb *qa:s-* ‘to tear, to be torn’ and a related transitive verb *qasə-* ‘to tear, to rip off’ (Watters 2006:149). The adjectival forms of the two verbs would mean ‘torn, torn off (something that is torn off)’ for the intransitive verb and ‘torn, torn off (something that has been torn off by somebody)’ for the transitive verb. A further semantic evolution ‘torn off’ > ‘separated, separate’ > ‘lonely, single’ > ‘one’ is not that difficult to conceive. Although no example of the adjectival forms meaning ‘torn, torn off’ is found in Watters (2006), they may be constructed as *qa:s-n* and *qa:sə-di* for the intransitive verb and *qas-n* and *qasə-di* for the transitive verb. Note, however, that “the *-di* form [...] is an adjectival form that is available to only a few transitive verbs”, i.e. it is typically used with intransitive verbs (Watters 2006:116-117). This may imply that instead of the four adjectival forms, we should rather expect just three, two intransitive ones, *V-di* and *V-n*, and one transitive *V-n*. An epenthetic *ə* in the form *qa:sə-di* is to be expected because no instances of the sequence *sd* have been found in the source, which suggests that such a sequence is unlikely to be tolerated by the language. Besides the epenthesis of a schwa, the assimilation of *sd* into *st* is also conceivable, which will give us *qa:s-ti* and *qas-ti* respectively. To this it may be added that the long vowel in the stem *qa:s-* is rather exceptional, since the source does not provide any other pair of verbs, an intransitive and a transitive one, based on the same root where the two forms would differ in their vowel length. Therefore, the vowel length in the intransitive stem may be of a secondary origin, the original forms of the two verbs being thus **qas-(t)* for the intransitive one and **qas-ə* for the transitive one, where *-t* and *-ə* are intransitive (or “anticausative”) and transitive (or “causative”) verbal markers respectively (on the latter two suffixes, cf. footnote 57 in this section).

That *nə(t)n* is interpreted as a non-human interrogative ‘what?’ and that *nəti* ~ *nədi* tends to be used as a human interrogative ‘who?’ can be explained as follows. First, the inclination to the differentiation between ‘who?’ and ‘what?’ as such may be due to the influence of the neighbouring languages, all of which do make this distinction. Recall in this respect that Kusunda is moribund and its speakers mostly use a different language in their everyday lives. Second, in Kusunda there exists a certain interplay between the (in)transitivity of the verb used to form a given adjectivized/nominalized verbal form and the (in)animacy of the head noun it modifies, which may have played a certain role as well.⁵⁸

intransitive respectively. The vowel *u* in *p^hiru-t-* is likely to be epenthetic in origin because the sequence *rt* is not attested elsewhere in the source, while the sequences *ua* ~ *wa* or *rw* are possible.

⁵⁸ At least, it might have been relevant for determining which of the forms would be chosen for the expression of the human interrogative ‘who?’ and which one for the non-human

Thus, in Kusunda the use of a deverbal adjective based on an intransitive verb implies that the action conveyed by the verb is presented as agentless, as happening, as it were, “on its own”, i.e. the head noun is an actor semantically, and thus necessarily animate (Watters 2006:116) as in (66a) vs. (66b), or at least an undergoer, as in (67a). On the contrary, the use of a deverbal adjective based on a transitive verb usually implies the existence of an agent and that the head noun is explicitly conceived as the patient of the action conveyed by the verb (see Watters 2006:115-116), as in (67b). At the same time, the two adjectivizers/nominalizers *-n* and *-di* differ in that the former is used with both transitive (67b) and intransitive verbs (66a, 67a), while the latter is typically used with intransitive verbs (68). However, it is remarkable that even when the verb is transitive, it seems that the *-di* form tends to exclude the patient reading, as in (69), which can be compared to (67b), and (70), where despite the verb being transitive, the nominalized form does not refer to the things stolen but to the person who stole them. Given the usual correlation between a higher agentivity and animacy/humanness, it is not surprising then that the interrogative *nəti* ~ *nədi*, presumably derived by means of the adjectivizer/nominalizer *-di*, would tend to be used as the human interrogative ‘who?’. In turn, the other interrogative, *nə(t)n*, presumably derived by means of the adjectivizer/nominalizer *-n*, has come to be restricted to the non-human meaning ‘what?’.

3.5 Tungusic languages: Evenki

Small Tungusic speaking communities are scattered over a vast territory including Eastern Siberia, the Russian Far East, and the neighbouring regions of China and Mongolia. Most Tungusic languages are endangered or moribund. According to a widespread although not universally accepted view, Tungusic languages (also known as Manchu-Tungus) are part of a larger Altaic phylum, which traditionally also includes at least Turkic and Mongolian.⁵⁹ Several internal classifications of the Tungusic languages have been proposed. Figure 1 illustrates the classification found in the Ethnologue. In other classifications, the Southeast subbranch may either be set apart as a third, “Central” branch (cf. Doerfer 1978) or grouped together with the Northern Tungusic languages (cf. Sunik 1968).

interrogative ‘what?’.

⁵⁹ This point of view is also adapted in the Ethnologue.

*Figure 1. Tungusic languages according to the Ethnologue***Northern Tungusic**

- Even
- Evenki, Oroqen
- Negidal

Southern Tungusic

- Southeast
 - Nanai, Oroq, Ulch
 - Udihe, Oroch
- Southwest
 - Jurchen
 - Manchu
 - Xibe

Interrogative pronominals used indiscriminately as ‘who?’ or ‘what?’ appear to exist only in some varieties of Evenki. Thus, Nedjalkov (1997) reports the ‘what?, who?’ interrogatives *e:-* and *e:-kun* for a variety of Evenki spoken in the Vanavara, Mutorai and Strelka-Chunja villages situated in the southern part of the Evenki National District. This variety, further conveniently called Vanavara Evenki, belongs to the so-called Southern Evenki group of dialects. Another ‘who?, what?’ interrogative *aji:* is reported by Bulatova & Grenoble (1999). The ‘who?, what?’ usage of the interrogatives *e:-* and *e:-kun* in Vanavara Evenki is probably due to interference with Russian, while the ‘who?, what?’ usage of *aji:* may result from its former use as a placeholder, somewhat similar to the English *whatchamacallit* or *whatsi(t)sname*. In what follows, I will first discuss the interrogative *aji:* in Section III.3.5.1 and then proceed to the discussion of the interrogatives *e:-* and *e:-kun* in Section III.3.5.2.

3.5.1 The interrogative aji:

According to Bulatova & Grenoble (1999:24), Evenki has the following three interrogative pronominals: two dedicated interrogatives, *ɲi:* ‘who?’ (71) and *e:kun* ‘what?’ (72), and one multifunctional interrogative *aji:*, which is “used with reference both to humans and inanimate objects” and which “fulfils the function of interrogative pronouns such as ‘who’, ‘what’, ‘which’, ‘how many’, ‘where’” (73).⁶⁰

⁶⁰ Given that Bulatova & Grenoble (1999) do not further precise the dialectal distribution of these interrogatives, I presume that these or similar forms are found in all or the majority of dialects.

Evenki

- (71) *ŋi:-du: kniga-βa bu:-rə-Ø?*
 who-DAT book-ACC give-AOR-3PL
 ‘Who did they give the book to?’ (Bulatova & Grenoble 1999:24)
- (72) *e:kun-ma ičə-čə:-s?*
 what-ACC see-PST-2SG
 ‘What did you see?’ (Bulatova & Grenoble 1999:24)
- (73) *aŋi:-βa nuŋan bu:-rə-n?*
 IPW-ACC 3SG[NOM] give-AOR-3SG
 ‘What/who/how many/what kind did he give?’ (Bulatova & Grenoble 1999:25)

The interrogative *aŋi:-* can also be used as a verb (74).⁶¹

Evenki

- (74) *aŋi: aŋi:-βa aŋi:-žə-ra-n?*
 IPW[NOM] IPW-ACC IPW-IPFV-PRS-3SG
 ‘Who is doing what?’ (Bulatova & Grenoble 1999:26)

According to Bulatova & Grenoble (1999:26) the use of *aŋi:-* as a verb is “widespread in some Evenki dialects, including those of the Podkamennaya Tunguska group [part of the Southern Evenki dialect group], which constitute the basis of the literary language”.

Interestingly, *aŋi:-* can also be used in non-interrogative utterances, where “it takes on a deictic function and can be translated as ‘this’, acquiring the role of a substantivized pronoun, 3rd person, singular or plural” (Bulatova & Grenoble 1999:25). The only example of this use provided by (Bulatova & Grenoble 1999:24), *aŋ bəjə* ‘this person’, is quoted from Cincius *et al.*’s (1975-1977:I:45) comparative Tungusic dictionary. Like the interrogative *aŋi:-*, the non-interrogative *aŋi:-* can also be used as a verb (75).

Evenki

- (75) *nuŋan aja-t aŋi:-du: aŋi:-žə-ra-n*
 3SG[NOM] good-ADV taiga-DAT PLH-IPFV-PRS-3SG
 ‘S/he lives well in the taiga’ (Bulatova & Grenoble 1999:26)

⁶¹ As such, this use is not so surprising taking into account the fact that a “weak distinction between the word class of roots” is typical for Tungusic languages (Whaley 1998). What is remarkable, however, is that in (74) *aŋi:-* ‘do what?’ may still take a complement *aŋi:-βa* ‘what-ACC?’, which appears to be quite exceptional for languages with interrogative pro-verbs (cf. Idiatov 2006).

The meaning of *aŋi:* (and the like) in Evenki may differ depending on the dialect. Furthermore, *aŋi:* (and the like) words can be found in some other Tungusic languages as well. The comparative data on *aŋi:* in various Tungusic languages provided in Cincius *et al.*'s (1975-1977:I:45) are reproduced here in (76).

(76) *The aŋi:-related forms in Tungusic languages (adapted from Cincius et al. 1975-1977:I:45)*⁶²

a. **Evenki** (Northern Tungusic)

aŋ ~ *aŋɛ* ~ *aŋi* (Southern Evenki dialects of the Podkamennaya Tunguska region, Tokma dialect of Southern Evenki, Agata and Bol'shoy Porog dialects of Northern Evenki, Urmi and Uchur dialects of Eastern Evenki):

1. ("interjection") 'yes, OK, all right, well';
2. 'yes, that very (one); this', e.g. *aŋi bəjə* 'this person'

aŋɛ ~ *aŋi* (Southern Evenki dialects of the Podkamennaya Tunguska Northern Baikal regions), *aŋi* (Zeya dialect of Eastern Evenki):

1. 'what?';
2. ("interjection") 'whatchamacallit?, whatsi(t)sname?';
3. (only in Zeya dialect of Eastern Evenki) 'really?, is it possible?'

b. **Negidal** (Northern Tungusic)

aŋ ~ *aŋɛ* ~ *aŋi:* (both Upper and Lower Amgun' dialects):⁶³
"interjection" 'yes, OK, all right, well'

c. **Udihe** (Southern Tungusic)

aŋi (Anjuj and Khor dialects): ("interjection") 'whatchamacallit?, whatsi(t)sname?' ("when trying to remember")

d. **Orok** (Southern Tungusic)

aŋŋv (Anjuj and Khor dialects): ("interjection") 'whatchamacallit?, whatsi(t)sname?' ("when trying to remember")

e. **Nanai** (Southern Tungusic)

aŋqa (Kur-Urmi dialect): 'properly, duly, thoroughly'

The comparative data summarized in (76) seems to suggest that the use of *aŋi:* as a full-fledged interrogative, as described for (some varieties of)⁶⁴ Evenki by

⁶² Translation from Russian and transliteration from Cyrillic is mine.

⁶³ The two dialects are also known as Verkhovskiy and Nizovskiy respectively.

⁶⁴ Bulatova & Grenoble (1999) do not specify the dialects where *aŋi:* can be used as an interrogative. Bulatova herself is a native speaker of an Eastern Evenki dialect of the Amur River region (Igor Nedjalkov, p.c.).

Bulatova & Grenoble's (1999), has originated in its use as a so-called universal placeholder or filler, as in (76a, 76c, 76d), somewhat similar to the English *whatchamacallit* or *whatsi(t)sname*. It is not so uncommon cross-linguistically that a placeholder can be used both for things *and* humans. For instance, in French *chose* 'thing' or *machin* from *machine* 'machine' can be used as placeholders both for things and humans. In Flemish, *dinge(s)*, derived from *ding* 'thing', is used somewhat similarly. Cf. also Section III.2.3.4.2.1 on the Old Arabic and modern dialectal *šay*⁶⁵ 'thing, something, somebody'.

Interestingly, like the French, Flemish or Arabic forms just mentioned, the Tungusic form *aŋi:* may also be related to the word meaning 'thing'. Compare, in this respect, the Evenki word *anŋa:* (Nepa dialect of Southern Evenki, Aldan dialect of Eastern Evenki) or *aŋa* ~ *aŋŋa* (Southern Evenki dialects of the Northern Baikal region) 'thing, object, item, utensil' (Cincius *et al.*'s 1975-1977:I:43). The final *-i:* of *aŋi:* would then probably come from the so-called "old genitive"⁶⁵ or "alienable possession marker" *-ŋi(:)*⁶⁶ (Bulatova & Grenoble 1999:13-15; Nedjalkov 1997:158-159) as in (77a) and (77b) respectively. The form *aŋi:* would then be structurally identical to the Flemish *dinges* or the similar German word *Dings*, which have also originated as genitives of the words meaning 'thing'.

Evenki

- (77) a. *Ivul-ŋi* *oro-r-in*
 PROP-GEN reindeer-PL-3SG.POSS
 'Ivul's reindeer (PL)' (Nedjalkov 1997:158)
- b. *bejumimni* *ulle-ŋi-n*
 hunter meat-APOSS-3SG.POSS
 'the hunter's meat' (Nedjalkov 1997:159)

Alternatively, *aŋi:* may be the fossilized genitive of the same 'what?' root as can be found, for instance, in the Vanavara Evenki interrogative *a-va-dy* 'which? (IPW-ACC-ADJZ)' (Nedjalkov 1997:214).⁶⁷

⁶⁵ This genitive is called "old" because "the possessive construction with the 'old genitive' was preferable some fifty years ago [...] but nowadays people use [...] possessive phrases mostly without the suffix *-ŋi*" (Nedjalkov 1997:158).

⁶⁶ According to Bulatova & Grenoble (1999:14) the "old genitive", which they call the "direct possessive suffix", is *-ŋi:*, while the alienable possession marker is *-ŋi*.

⁶⁷ Cf. Cincius *et al.* (1975-1977:1:4-5) for other similar (non-human) *a*-initial interrogatives in Evenki and other Tungusic languages. In fact, the initial *a-* of the Vanavara Evenki *avady* 'which?' is in all probability also cognate to the Vanavara Evenki interrogative *e:-* 'what?' (cf. Section III.3.5.2).

3.5.2 The interrogatives *e:-* and *e:-kun*

According to Nedjalkov (1997:7-11, 214-215), Vanavara Evenki uses a dedicated human interrogative pronominal *ŋi* ‘who?’ (78) and two ‘what?, who?’ interrogatives *e:-* (79-81) and *e(:)-kun* (82-85), where *-kun* is said to be an “intensifier”. The interrogative *e:-* can also be used as a verbal root ‘do what?’ (81), similarly to *aŋi:* discussed in Section III.3.5.1.

Vanavara Evenki

- (78) *ŋi ta-du ŋene-žere-n?*
IPW[NOM] that-DAT go-PRS-3SG
‘Who is going there?’ (Nedjalkov 1997:10)
- (79) *e:-va genno-sin-čə-s?*
IPW-ACC bring-SMLF-PST-2SG
‘What did you go to fetch?’ (Nedjalkov 1997:10)
- (80) *nuŋan Hovoko-nun ta-duk e:-nun evi-žere-n?*
3SG[NOM] PROP-COM that-ABL IPW-COM play-PRS-3SG
‘He is playing with Hovoko, and with who else?’ (Nedjalkov 1997:11)
- (81) *e:-ra-n tatkit-tu?*
IPW-NON<FUT>-3SG school-DAT
‘What did s/he do at school?’ (Nedjalkov 1997:8)
- (82) *si e:kun-ma e-du o:-žə-nni?*
2SG IPW-ACC this-DAT do-PRS-2SG
‘What are you doing here?’ (Nedjalkov 1997:9)
- (83) a. *e:kun tari (bi-si-n)?*
IPW[NOM] that[NOM] be-PRS-3SG
‘Who/what is that/this?’ (Nedjalkov 1997:9)
- b. *tar ekun? nuŋan bi-si-n*
IPW[NOM] that[NOM] 3SG be-PRS-3SG
‘Who is that? That’s he’ (Nedjalkov 1997:196)
- (84) *e:kun-duk?*
IPW-ABL
‘from whom/where?’ (Nedjalkov 1997:9)
- (85) *tar oron e:kun-ŋi?*
that reindeer IPW-GEN
‘Whose is that reindeer?’ (Nedjalkov 1997:10)

The derived form *e:-kun* is more common than the simple *e:-*. Furthermore, only *e:-kun* can occur as an independent word in the (unmarked) nominative, i.e. *e:-* is a bound root and does not have a nominative form.

As can be readily observed in (79-85), whether *e:-* and *e(:)kun* are interpreted as a human ‘who?’ or a non-human ‘what?’ very much depends on both the larger context and the immediate morphological marking of these interrogatives, such as the case marker they are used with. Thus, the use of the comitative form *e:-nun* in (80) or the genitive *e:kun-ŋi* in (85) triggers the human interpretation, because the genitive *-ŋi* is typically used for “personal possession” (Nedjalkov 1997:158), while “the comitant in *-nun* always denotes an animate participant with the same semantic and pragmatic status as the participant-‘syntactic possessor’” (Nedjalkov 1997:155).⁶⁸ In (79) and (82), the non-human interpretation of *e:-* and *e:kun* as ‘what?’ is strongly favoured both by the larger context and the accusative case they are used in.⁶⁹ In (84), the use of the ablative suffix *-duk* allows both for a human and a non-human interpretation of the interrogative, although the non-human interpretation would probably be more common. In (83a), both a human and a non-human interpretation of the nominative form *e:kun* appear to be possible, while in (83b) the larger context suggests a human interpretation of the same form. The presence of verbal morphology on *e:-* in (81) favours a non-human reading.

The former meaning of the interrogative root *e:-* and its derivative *e(:)kun* ‘who?, what?’ must have been only ‘what?’. This is supported both by comparative data from other Tungusic languages, including some closely related varieties of Evenki, and by some internal evidence. For instance, a piece of internal evidence is provided by the Vanavara Evenki specific indefinite pronominal *ekun-mal*, whose root is identical to the interrogative at issue and which “usually indicat[es] an inanimate object”, even though it “may also indicate an animate referent [...] even if the referent is known to be human” (Nedjalkov 1997:198), (86).

Vanavara Evenki

- (86) *ekun-mal iče-v-žere-n?*
 INDF[NOM] see-PASS-PRS-3SG
 ‘Something/somebody is seen’ (Nedjalkov 1997:198)

Abundant comparative evidence for the former non-human meaning ‘what?’ of the Vanavara Evenki interrogative *e:-/e(:)kun* can be found in Cincius *et al.* (1975-1977:I:3-5, 286-288). Thus, the cognates of the interrogative root *e:-*, as *e:-*

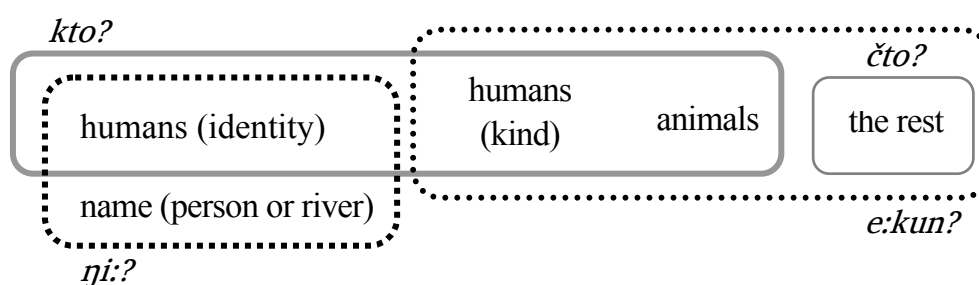
⁶⁸ It should be mentioned, however, that the comitative is not always considered as case in descriptions of Evenki. Thus, unlike Bulatova & Grenoble (1999:8), Nedjalkov (1997:142) does not include the comitative in the case paradigm of Evenki nominals.

⁶⁹ It is often remarked that direct objects are normally less agentive than subjects and are prototypically non-human or inanimate.

/æ:-/i:-/ja-/(x)aj-/(x)ai-/a- and the like, are found in all Tungusic languages and only in the meaning ‘what?’.⁷⁰

Before proceeding further with the discussion of the origins of the Vanavara Evenki ‘who?, what?’ interrogative, a caveat needs to be made. In descriptions of Tungusic languages, *e:-* and cognate interrogatives are often found to be glossed as ‘who?, what?’. However, this is just due to the fact that these descriptions are made in Russian. In Russian, the interrogative *kto* ‘who?’ will be used in several contexts where many Tungusic languages would use *e:-* or a cognate ‘what?’ interrogative, in fact very much like English *what?*. By way of illustration, Figure 2 provides a comparison between the Poligus Evenki interrogative pronominals *ŋi:* ‘who?’ and *e:kun* ‘what?’, as described in Konstantinova (1968), and their Russian counterparts, *kto* ‘who?’ and *čto* ‘what?’.

Figure 2. Russian *kto* ‘who?’ and *čto* ‘what?’ vs. Poligus Evenki *ŋi:* ‘who?’ and *e:kun* ‘what?’ (based on Konstantinova 1968:72-73)



Example (87) illustrates the use of *e:kun* in a KIND-question in Poligus Evenki.⁷¹

Poligus Evenki (Konstantinova 1968:73)

- (87) *e:kun bi-si-nni? bi Kurkogir bi-si-m*
 what be-PRS-2SG 1SG PROP be-PRS-1SG
 ‘[A:] What are you? [B:] I am a Kurkogir (i.e., I belong to the Kurkogir clan, the name of my family is Kurkogir)’

Example (88) illustrates the use of the Poligus Evenki *ŋi:* ‘who?’ in a question about the name of a river.

⁷⁰ It may also be interesting to mention that in their comparative Tungusic dictionary under the entry *abgū* ‘which?’ summarizing *(x)aj-/(x)ai-/a-* and the like Tungusic interrogatives and their derivatives, Cincius *et al.* (1975-1977:3-5) reconstruct it as **kaj* for Proto Tungusic and compare it to the Old Turkic *qaju* ‘which [N]?, what [N]’, *qajda* ‘where?’ and the Buryat (Eastern Mongolian) *xaa* ‘(to) where?’, *xajša* ‘from where?’.

⁷¹ Cf. Konstantinova (1968:73): “*e:kun* is also used in questions about unknown persons (if the question is about their profession, sex, clan)”.

Poligus Evenki (Konstantinova 1968:73)

- (88) *ær bi:ra ŋi: gərbi:-n?*
 this river[NOM] who name-3SG.POSS
 ‘What (lit.: ‘who?’) is the name of this river?’

Interestingly, the use of *ŋi:* ‘who?’ or another cognate form in the context of a non-prototypical combination of values of the kind [thing + identification + proper name] (cf. Section II.3), as in (88), normally does not result in the gloss ‘who?, what?’ in the sources. In all probability, the reason is that Russian would use neither *kto* ‘who?’ nor *čto* ‘what?’ here, but rather *kak* ‘how?’ (89a-c) or *kakoj* ‘what [N]?’ (89d).

Russian

- (89) a. *kak ego imja?*
 how 3SG.M.GEN name.N.SG.NOM
 ‘What is his name? (a person)’
 b. *kak ego zovut?*
 how 3SG.M.GEN call.IPFV.PRS.3PL
 ‘What is he called? (a person)’
 c. *kak nazyvaetsja eta reka?*
 how call.IPFV.PRS.3SG.REFL this.F.SG.NOM river.F.SG.NOM
 ‘What is this river called?’
 d. *kakoe nazvanie (u) etoj reki?*
 what.N.SG.NOM_{ADJ} appellation at this.F.SG.GEN river.F.SG.GEN
 ‘What is the name of this river? (lit.: ‘What name is of this river?’ or
 ‘What name does this river have?’)

Unlike Russian, all other Tungusic languages are like Poligus Evenki in using their ‘what?’ interrogative in questions about animals. Besides Evenki, the use of ‘what?’ in KIND-questions can, for instance, be found in Udihe (90-92) and Nanai (93). Bystrinskiy, Penzhinskiy and Olskiy dialects of the Eastern Evenki dialect group are also reported to have an interrogative verb *æ:-t-* ‘be what to somebody? (in the same sense as the Udihe example (92))’ (Cincius *et al.* 1975-1977:I:287), where *æ:-* is cognate to the Evenki *e:-* ‘what?’ and *-t* looks like a frozen instrumental suffix.

Udihe (Southern Tungusic)

- (90) *si mamase-i j’e-we b’a:-ni? aziga-wa-si*
 2SG wife-2SG.POSS what-ACC give.birth.PST-3SG girl-ACC-or
b’ata-wa-si?
 boy-ACC-or
 ‘What did your wife give birth to, a girl or a boy?’ (Maria Tolskaya, p.c.)

- (91) *nuani j'eu bi:-ni? niŋka*
 3SG IPW be-3SG.PRS Chinese
 '[A:] What is s/he? [B:] A Chinese' (Maria Tolskaya, p.c.)
- (92) *si Iwana-zi i:-mule bi-u za:-mule-s xa:-mule-s?*
 2SG PROP-INS what-AFF be-2PL cousin-AFF-OR brother-AFF-OR
 'What are you and Ivan to one another, cousins or brothers?' (Nikolaeva & Tolskaya 2001:804)

Torgon Nanai

- (93) *ej arčoka:n si xaj-si?*
 this girl 2SG what-2SG.POSS
 'What is this girl to you?' (Onenko 1980:440)

Besides Evenki, the use of 'who?' in the context of a non-prototypical combination of values of the kind [thing + identification + proper name] (only personal proper names) can be found in Even (94) (Aleksey Burykin, p.c.), Negidal (Cincius *et al.* 1975:I:660), Orok (Cincius *et al.* 1975:I:660), and Manchu (Avrorin 2000:113), but not in Nanai (Onenko 1980) or Udihe (Maria Tolskaya, p.c.), for instance.

Even

- (94) *gerbe-s ni?*
 name-2SG.POSS who
 'What is your name?' (Aleksey Burykin, p.c.)

Let us now turn back to the discussion of the origins of the Vanavara Evenki 'who?', 'what?' interrogative *e:-* (and *e(:)kun*). First of all, note that Vanavara Evenki is very closely related to Poligus Evenki, which was used in Figure 2 to illustrate the typical uses of the 'what?' interrogative pronominals in Tungusic languages. Both dialects belong to the Podkamennaya Tunguska subgroup of the Southern Evenki dialect group. Still, the Poligus Evenki interrogative pronominal *e:kun* is a regular non-human interrogative 'what?', which can also be used in the case of a non-prototypical combination of values of the kind [person + classification (+ common noun)] (cf. Section II.2), and *ŋi:* is a regular human interrogative 'who?'. The two interrogatives are very much like English *what?* and *who?*, as has been illustrated in Figure 2. One might therefore suspect that the Vanavara Evenki *e:-* and *e(:)kun* are simply functionally identical to the Poligus Evenki *e:kun* and therefore, are not real 'who?', 'what?' interrogatives. However, some examples provided by Nedjalkov (1997), such as (80) and (85), hardly lend themselves to such an interpretation and rather suggest that Vanavara Evenki *e:-* and *e(:)kun* are indeed the 'who?', 'what?' interrogatives we are looking for.

The shift from ‘what?’ to ‘who?, what?’ in the use of *e:-* and *e(:)kun* in Vanavara Evenki should probably be explained by an interference with Russian. Recall, in this respect, that Evenki is an endangered language with (almost) all speakers being bilingual in Russian and very few Evenki mother tongue speakers in younger generations. Given that in Russian *kto* ‘who?’ will be used in several contexts where Evenki would normally use *e:-* or *e(:)kun*, cf. Figure 2 above, Vanavara Evenki speakers might have reinterpreted the original ‘what?’ interrogatives *e:-* or *e(:)kun* as an equivalent not only to the Russian *čto* ‘what?’, but also *kto* ‘who?’ and not only in questions about animals and in KIND-questions, but also, apparently by extension, in regular questions about the identity of persons.

4 Southeast Asia and Oceania

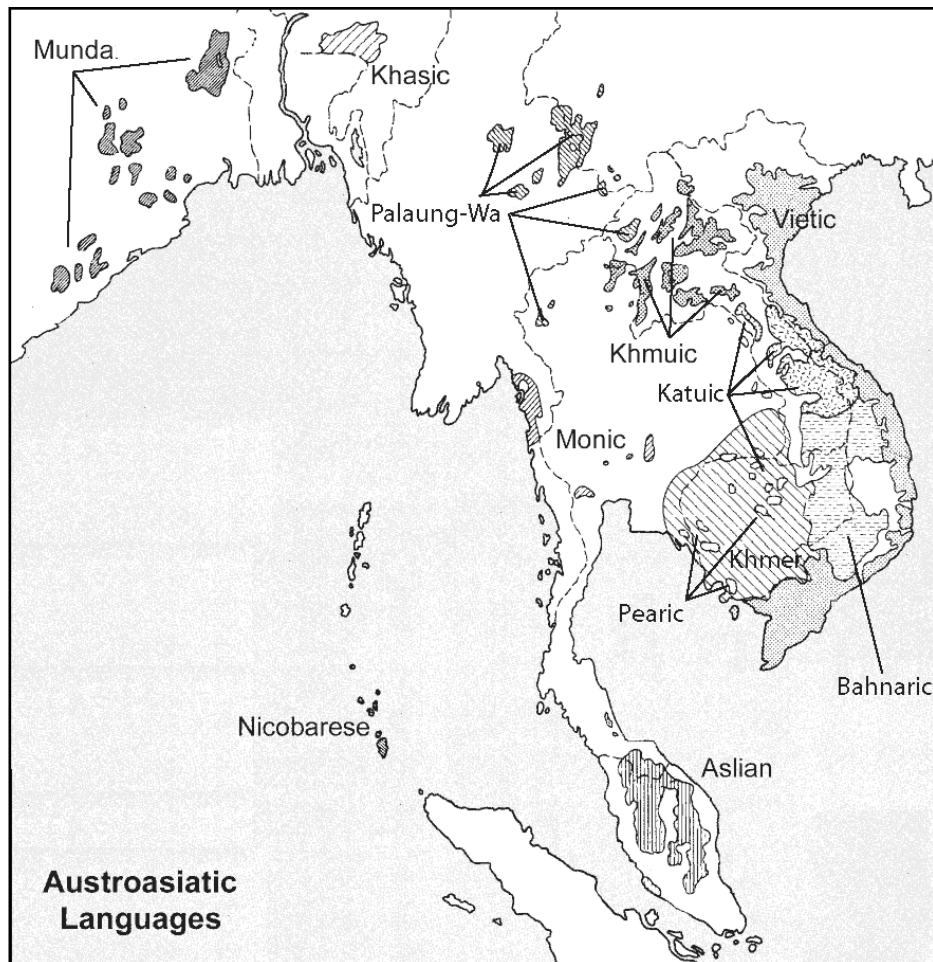
In this section, I will discuss the languages presumably allowing a lack of differentiation between ‘who?’ and ‘what?’ that are spoken in Southeast Asia and Oceania. The section is organized as follows. In Section III.4.1, I will examine Austro-Asiatic languages and in Section III.4.2 Austronesian languages.

4.1 Austro-Asiatic languages

4.1.1 Introduction

Austro-Asiatic languages are spread through a vast area including continental South-East Asia, Eastern India and the Nicobar Islands, Map 1.

Map 1. Distribution of Austro-Asiatic languages
(<http://www.anu.edu.au/~u9907217/languages/languages.html>)¹



¹ This map does not represent some smaller Austro-Asiatic languages spoken in China, e.g. Bolyu/Palyu/Lai, Bagan, etc. Furthermore, a Munda language Korku is spoken in India further to the west of the area covered by this map.

The received view is that the Austro-Asiatic phylum is subdivided in two major blocks, Munda languages in the west and Mon-Khmer languages in the east. A further classification as for instance adopted in the *Ethnologue* is summarized in Figure 1.

Figure 1. Classification of Austro-Asiatic languages according to the Ethnologue

Mon-Khmer

Northern Mon-Khmer: Khasian, Palaungic, Khmuic, Mang

Palyu

Eastern Mon-Khmer: Bahnaric, Katuic, Khmer, Pearic

Viet-Muong

Monic

Aslian

Nicobar

Unclassified: Bugan, Buxinhua, Kemiehua, Kuanhua

Munda

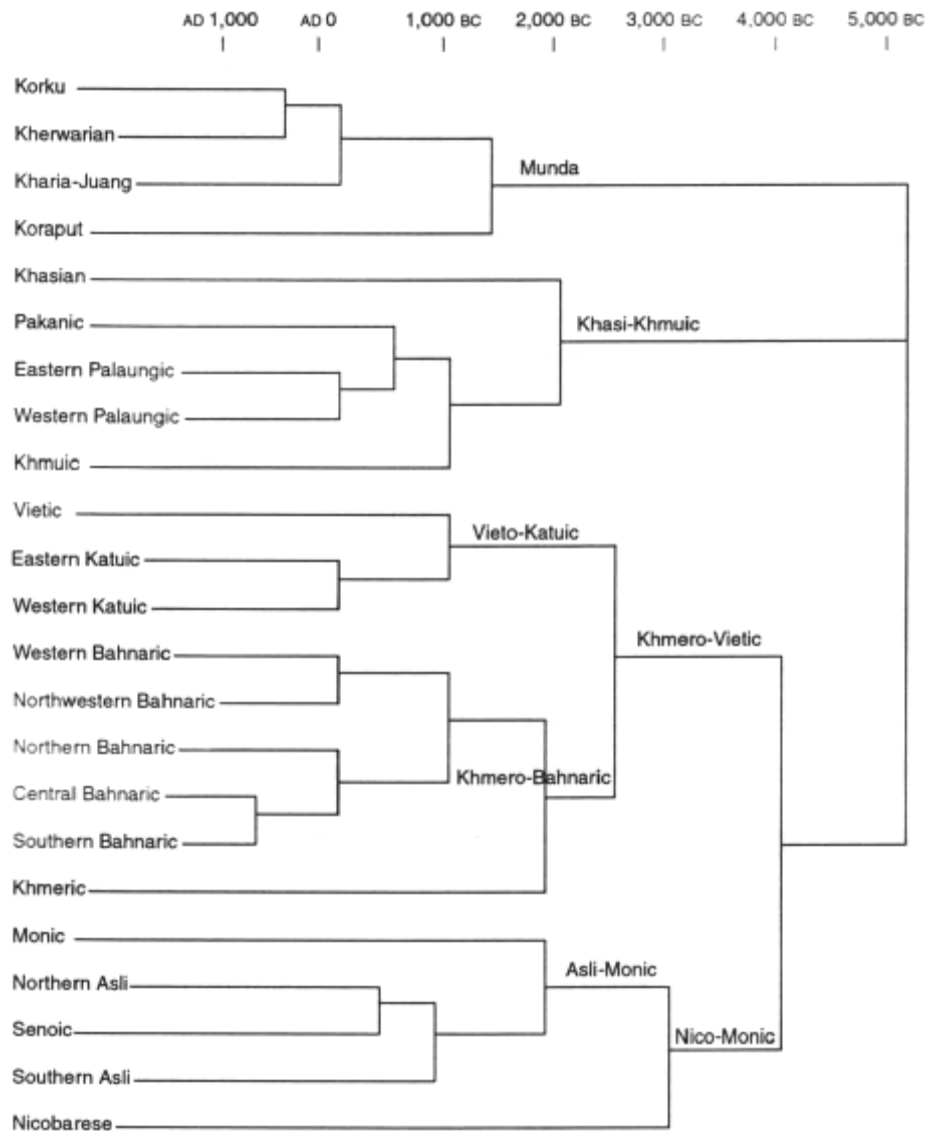
North Munda: Kherwari, Korku

South Munda: Kharia-Juang, Koraput Munda

However, there is no real consensus about the general bipartite division or the number of subgroups of “Mon-Khmer”, which in certain classifications may well exceed ten (cf. Sidwell 2005, 2006 for a brief overview of the existing classifications). For instance, Diffloth (2005) proposes a tripartite classification with Northern Mon-Khmer languages together with Palyu and Bugar being considered as a third major branch of Austro-Asiatic, “Khasi-Khmuic”, on a par with Munda and the rest of the traditional Mon-Khmer family, Figure 2. As I am not in position to make educated judgements on various classifications, I will arbitrarily adopt the classification proposed in the *Ethnologue* for the purpose of reference.

Languages with interrogative pronominals used indiscriminately as ‘who?’ or ‘what?’ appear to exist both among Munda and Mon-Khmer languages, and in both Northern Mon-Khmer (or Khasi-Khmuic in Diffloth’s 2005 terms) and non-Northern Mon-Khmer languages: Khasian (Section III.4.1.2), (Eastern) Palaungic (Section III.4.1.3), Bahnaric (Section III.4.1.4), Aslian (Section III.4.1.5), and Munda (Section III.4.1.6). Judging from the data available, each of the branches at issue appears to have just a few such languages, with the remaining majority of languages distinguishing between ‘who?’ and ‘what?’. The only noticeable excep-

Figure 2. Classification of Austro-Asiatic languages according to Diffloth (2005)



80tion may be Khasian, but this is due to the fact that it is a branch with relatively little internal diversification.²

Interestingly, although formally the ‘who?, what?’ interrogatives in various Austro-Asiatic languages are rather different, the ways these interrogatives have come to existence appear to be strikingly similar in language after language. Thus, it appears that in all non-Munda Austro-Asiatic languages the ‘who?, what?’ interrogative has developed out of an interrogative literally meaning ‘what one?’, i.e. ‘which one? (person or thing)’. This phrasal interrogative must have been constructed of two elements: (i) a deictic or pronominal root, which in

² In fact, until relatively recently, Khasian used to be considered a single language branch.

at least one case is the numeral ‘one’, and which is indifferent to the human vs. non-human distinction, (ii) a postposed interrogative ‘what?’ used as a modifier, which commonly seems to have been a reflex of a Proto Mon-Khmer (i.e., non-Munda) **m[o]ʔ/*m[o]h* ‘what?’ (Shorto 2006, #136).³ Later, in some cases, the phrasal selective interrogative has been generalized to non-selective contexts for questions on both persons and things, while in the others this generalization has happened along with a reduction of the original interrogative modifier, similar to the way Italian *cosa?* ‘what?’ has developed out of *che cosa?* ‘what thing?’.

Note that although the development of the selective ‘which one? (person or thing)’ (literally ‘what one?’) into the non-selective ‘who?, what?’ has happened here and there, its development to the non-human ‘what?’ appears to have been much more common. In all probability, this was due to the literal meaning ‘*what one?*’.

The development of ‘who?, what?’ interrogatives in Munda is actually quite similar to what is suggested for the non-Munda Austro-Asiatic languages. The main difference seems to lie in a slightly different starting point from which Juang, a Munda language presumably allowing for a lack of differentiation between ‘who?’ and ‘what?’, has departed, but further semantic shifts seem to follow the already familiar pattern. Thus, the original interrogative involved seems to have been locative ‘where?’, which later has developed into a selective interrogative pronominal ‘which one? (person or thing)’ and has ultimately become extended to non-selective contexts both in questions about persons and things.

In what follows, I will further elaborate on the Austro-Asiatic languages with ‘who?, what?’ interrogatives. To avoid the complications related to the issue of the internal classification of the Austro-Asiatic phylum, the discussion will be organized according to the universally excepted lower-level subgroups: Khasian (Section III.4.1.2), (Eastern) Palaungic (Section III.4.1.3), Bahnaric (Section III.4.1.4), Aslian (Section III.4.1.5), and Munda (Section III.4.1.6).

³ Such phrasal way of deriving their interrogative pronominals (and often other kinds of interrogatives as well) appears to be very typical of non-Munda Austro-Asiatic languages in general. It is worth mentioning, in this respect, that the only interrogative pronominal that can be reconstructed more or less safely for the Proto non-Munda Austro-Asiatic level seems to be **m[o]ʔ/*m[o]h* ‘what?’ (Shorto 2006, #136), which in all probability could also be used attributively as ‘what [N]?, which [N]?’. Note that at the same time, no interrogative meaning ‘who?’ appears to be reconstructible to this level and in most cases the etymologies of ‘who?’ interrogatives in non-Munda Austro-Asiatic languages are quite transparent. Consider, for instance, the reconstructions involving the meaning ‘who?’ in Shorto (2006), especially their distribution.

4.1.2 Khasian languages

Khasi, spoken in India and Bangladesh, constitutes a group of its own within the Northern subdivision of the Mon-Khmer branch of the Austro-Asiatic phylum. There are at least four major Khasi varieties: the so-called Standard Khasi (or Khasi proper), War, Pnar (or Synteng) and Lyngngam. Traditionally, these varieties used to be considered as dialects, however in more recent accounts they are rather referred to as languages. Interrogative pronominals which can be used both as ‘who?’ and ‘what?’ can be found in at least three of the four Khasi idioms: Standard Khasi, War and probably Pnar. The fourth variety, Lyngngam, may have two distinct interrogatives ‘who?’ and ‘what?’, but the data are scarce and not always clear.

The existence of forms used indiscriminately as ‘who?’ or ‘what?’ in Khasi is due to the peculiarities of the Khasi gender systems, the fact that the Khasi interrogative pronominals are typically constructed of a gender-number marker preposed to a general interrogative pronominal root (*no* or *ey*), and that these gender-number markers can, or must be omitted in certain morphosyntactic contexts. Therefore, it will be useful to begin with a brief discussion of the Khasi gender systems in Section III.4.1.2.1 and then proceed to a general presentation of the Khasi interrogative pronominal systems in Section III.4.1.2.2. The emphasis will primarily be on Standard Khasi (Section III.4.1.2.2.1). Other varieties will be examined as far as is allowed by the data available (Section III.4.1.2.2.2). By way of conclusion (Section III.4.1.2.3), I will briefly discuss the possible origins of the general interrogative pronominal roots *no* and *ey*.

4.1.2.1 Khasi gender systems

Gender is present in all of the four major Khasi varieties. Gender is marked on every Khasi nominal together with number by means of a preposed gender-number marker, often referred to as an “article”.⁴ It should be mentioned that the gender-number markers may sometimes be omitted both on controllers and targets.⁵ In the singular all Khasi varieties distinguish at least two genders,

⁴ At least in Standard Khasi and Pnar, the gender-number markers are identical to the third person pronouns. With nouns and verbs the gender-number markers are usually graphically represented as separate words. In other contexts, e.g. with demonstrative or interrogative roots, they are usually written as prefixes. For the sake of uniformity, I write them separately throughout. No claims on their exact morphological status are implied hereby. I gloss them similarly to pronouns with verbs (e.g., *ka mut* = 3SG.F mean ‘it means’ in (3)) and as gender-number markers with other kinds of elements (e.g., *ka ta* = F.SG DEM ‘that’ in (3)).

⁵ The rules determining when the gender-number marker can or must be left out seem to be rather complex, with a good deal of lexical conditioning. The following statement in Grierson

masculine marked by *u*⁶ and feminine (or non-masculine) marked by *ka*. In the plural, which is marked by *ki*,⁷ no gender distinctions are made. All nouns designating male humans and male domestic animals are assigned to the masculine gender and those designating female humans and female domestic animals to the feminine gender. For all other nouns gender assignment is in general unpredictable (see Rabel 1961:93). Thus, in Standard Khasi *u riisaŋ* ‘squirrel’, *u dkhiw* ‘ant’, *u sntiw* ‘flower’ and *u ŋap* ‘bee’ are masculine, while *ka sier* ‘deer’, *ka wi*² ‘worm’, *ka iij* ‘house’ and *ka ŋap* ‘honey’ are feminine (Grierson 1906:II:8; Rabel 1961:93-94). Furthermore, in Standard Khasi, the feminine is the gender of abstract nouns and the feminine agreement pattern is used for the purposes of enforced agreement⁸ (cf. Grierson 1906:II:9; Rabel 1961:94-95; Rabel-Heymann 1977:247, 249), as in (1-3).

Standard Khasi (Austro-Asiatic, Northern Mon-Khmer, Khasian; India)

- (1) *ka dey ya ŋaa ba n loŋ*
 3SG.F be to 1SG REL FUT be
 ‘I ought to be (lit.: ‘It is necessary for me to be’)’ (Grierson 1906:II:10)
- (2) *ka la slap*
 3SG.F PST rain
 ‘It rained’ (Grierson 1906:II:9)
- (3) *ka ta ka mut*
 F.SG DEM 3SG.F mean
 ‘That means...’ (Rabel 1961:94)

In addition, Khasi may have one more gender, the diminutive (or diminutive-honorific), marked by *i*,⁹ which replaces the other gender-number markers. According to Rabel (1961:95), in Standard Khasi the diminutive is indifferent to number, although Grierson (1906:II:8) rather describes it as singular. The view of the diminutive as a gender on a par with the masculine and the feminine does not seem to enjoy a unanimous support among the scholars. Thus, Rabel (1961:95)

(1906:II:8) is quite illustrative in this respect: “The ‘article’ is omitted in idiomatic sentences when no ambiguity is caused by the omission”.

⁶ Sometimes, also written as ([?])*uu*. I use only the writing *u* for reference purposes, but I usually preserve the original orthography of the examples.

⁷ Sometimes, also written as *kii*. I use only the writing *ki* for reference purposes, but I usually preserve the original orthography of the examples.

⁸ The latter term refers to the use of agreement markers in cases when an atypical controller such as an interjection is used or no controller is possible at all. In a gender language, if an agreement target *can* agree then typically it *must* agree (Corbett 1991:204).

⁹ Sometimes, also written as ([?])*ii*. I use only the writing *i* for reference purposes, but I usually preserve the original orthography of the examples.

writes that “the article /ʔii/ and the corresponding personal pronoun /ʔii/ distinguish neither number nor gender”. Similarly, Corbett (1991:206) refers to Standard Khasi as a language with a “two-gender system”. That is, for these scholars Khasi has only two genders, the masculine and the feminine genders, with the diminutive-honorific simply blocking the formal expression of gender-number. In all probability, this analysis is due to the assumption that there are no nominal controllers that are lexically specified as diminutive-honorific. When applied restrictively, the latter assumption would imply that there should be no nominal roots which are marked exclusively with *i*. As far as I can judge from the data available, this is indeed what seems to be the case in Standard Khasi.

However, this restrictive analysis is not unproblematic. Thus, judging from the translations of pairs like *u briiw* ‘man’ (masculine) vs. *i briiw* ‘dwarf’ (diminutive) and *ka iij* ‘house’ (feminine) vs. *i iij* ‘hut’ (diminutive) provided by Grierson (1906:II:8), the relation between the diminutive noun and its non-diminutive counterpart may sometimes be non-compositional. Thus, a dwarf is not exactly the same as a small man and a hut is not really just a small house. This may be interpreted as evidence that diminutive nouns like *i briiw* ‘dwarf’ and *i iij* ‘hut’ are lexically specified as diminutive (at least in these very meanings, because probably *i briiw* can still be used to refer to a ‘small man’ and *i iij* to a ‘small house’).

Furthermore, according to Grierson (1906:II:17, 24, 30), in the non-Standard Khasi varieties such as Lyngngam, Pnar and War “the diminutive [...] is often used without any apparent reason, – possibly as a neuter”, e.g. Lyngngam *i rinon* ‘property’, Pnar *ha i tu i por* ‘at that time’ (lit.: at DIM DEM DIM time) and Pnar and War *i bhah* ‘portion, share’. That is, the diminutive marker seems to be often used in situations where Standard Khasi would use the feminine. In this respect, compare also the agreement patterns used for enforced agreement in non-Standard Khasi varieties in (4b-d) and (5b-d) to the feminine used for the same purpose in Standard Khasi in (4a) and (5a).¹⁰

Standard Khasi

- (4) a. *ka ba ŋaa kinmaaw šaphaŋ u jumaay*
 F.SG REL 1SG recollect concerning M.SG earthquake
 ‘What I recollect of the earthquake’ (the title of a text; Grierson 1906:II:15)

¹⁰ At certain places, the translations and the glosses of the examples from Grierson (1906:II) have been slightly modified.

Pnar

- b. *i wo u-kinmo ia u kh'mi*
 DIM REL INF-recollect of M.SG earthquake
 'What is to be recollected about the earthquake' (the title of a text;
 Grierson 1906:II:29)¹¹

War

- c. *i a ju-kimmo ia u kh'mi*
 DIM REL INF-recollect of M.SG earthquake
 'What is to be recollected about the earthquake' (the title of a text;
 Grierson 1906:II:36)

Lyngngam

- d. *ummet ne taa-ha-joŋ timma u jawmay*
 M.SG.?(REL) 1SG just-at-time recollect M.SG earthquake
 'What I recollect about the earthquake right now' (the title of a text;
 Grierson 1906:II:23)¹²

Standard Khasi

- (5) a. *kat ka ba u le?*
 so.much.as F.SG REL 3SG.M do
 'Whatever he does' (Nagaraja 1985:47)

Pnar

- b. *kat i ba em i joŋ ŋa, du i*
 so.much.as DIM REL be 3.DIM of 1SG only 3.DIM
joŋ mee don
 of 2SG be
 'Whatever I have is yours (as well)' (Grierson 1906:II:28)

War

- c. *kat i wa ah i joŋ ñia, i te*
 so.much.as DIM REL be 3.DIM of 1SG DIM DEM

¹¹ In the original the example is written as *i wow kənmo* [...], but the final *-w* here clearly does not belong to the relative pronoun but is the infinitive marker *u*, similar to the War infinitive marker *ju* in (4c).

¹² The exact gloss of *ummet* is somewhat unclear. The *-met* part is probably the same form as the interrogative *met* 'what?'. The *m* preceding *-met* may belong to the masculine marker, which would then appear here as *um-*. Otherwise, *m* may be a form of the possessive marker *ba/ am-ba/ amb/ am* 'of' (cf. (5d)), where *am* also means 'from' (Grierson 1906:II:17) and *ba* is identical to the relative pronominals of the other Khasi varieties (cf. (4a)).

i joŋ em
3.DIM of 2SG

‘Whatever I have is yours (as well)’ (Grierson 1906:II:34-35)

Lyngngam

d. *u met u met prok joŋ ne bad amba mi*
M.SG what M.SG what all of 1SG and of 2SG

‘Whatever I have is yours (as well)’ (Grierson 1906:II:22)

Note that where Pnar (4b, 5b) and War (4c, 5c) use the diminutive, Lyngngam uses the masculine (4d, 5d). Actually, even Standard Khasi sometimes seems to use the diminutive instead of the feminine for enforced agreement, as in (6). Although probably the appearance of examples like (6) is due to dialectal variation within Standard Khasi.

Standard Khasi

(6) *ʔii mat ba...*

3.DIM seem REL

‘It seems that...’ (Rabel 1961:133)

All in all, from a cross-Khasi perspective it seems preferable to treat the diminutive as a gender similar to, although not exactly on a par with the masculine and the feminine. Its gender status is clearly more solid in non-standard varieties of Khasi, such as Pnar, War and Lyngngam.

4.1.2.2 Khasi interrogative pronominal systems

Interrogative pronominals of the four major Khasi varieties are summarized in (7).

(7) *Interrogative pronominals of the four major Khasi varieties (based on Grierson 1906:II; Nagaraja 1985, 1996; Rabel 1961, Subbarao & Temsen 2003)*¹³

Standard Khasi (gender-number marker) + *no/ey* ‘who?, what?, which?’
ayu ‘what?’

¹³ The three interrogative roots of Standard Khasi are written somewhat differently in the sources. The writings presented in (7) come from Nagaraja (1985) and will be used elsewhere for citation purposes. Rabel (1961) prefers the transcriptions /nu/, /ʔey/ and /ʔayuu/ instead, while Grierson (1906:II) gives *no*, *ei* and *aiuh* and Subbarao & Temsen (2003) *no*, *Ei* and *ayu*. For the non-standard varieties the original orthographies as found in Grierson (1906:II) have been preserved.

War	(gender-number marker) + <i>ai</i> ‘who?, what?, which?’
Pnar	gender-number marker + <i>i</i> ‘who?, what?, which?’
Lyngngam	<i>yət</i> or <i>u</i> (M.SG) + (<i>i</i>) <i>et</i> ‘who?’ <i>mət/met</i> ‘what?’

In what follows, I will discuss these forms in more detail. Given the availability of data, the emphasis will primarily be on the Standard Khasi forms. The interrogative pronominal systems of War and Pnar seem to resemble that of Standard Khasi, while that of Lyngngam appears to be rather different.

4.1.2.2.1 Standard Khasi

Standard Khasi has two kinds of interrogative pronominals. The first kind is represented by the interrogative pronominals *no* and *ey*, both meaning ‘who?, what?, which?’ (Rabel 1961:67), which can, and sometimes must be marked by a preposed gender-number marker or by a preposition. The second kind is represented by the interrogative pronominal *ayu* ‘what?’, which is never marked by a preposed gender-number marker, neither can it be immediately preceded by a preposition. I will begin with a discussion of the interrogative pronominals *no* and *ey* (Section III.4.1.2.2.1.1) and then proceed to the interrogative *ayu* (Section III.4.1.2.2.1.2).

4.1.2.2.1.1 The interrogatives *no* and *ey*

In the present section it will be shown that in certain contexts *no* and *ey* may be opposed as selective vs. non-selective, but on the whole, in their use as interrogative pronominals this opposition is blurred. Neither in their selective nor in their non-selective uses does the difference between the two directly depend on the opposition human vs. non-human. Rather, *ey* is preferred in contexts where normally the interrogative must be marked for gender-number, whereas *no* is preferred in contexts where gender-number marking of the interrogative is not allowed. Whether in this or that particular case *no* and *ey* should be interpreted as ‘who?’ or ‘what?’ (or either) depends on two factors, (i) the overall context and/or (ii) the element preceding the interrogative and directly relating to it, such as a gender-number marker or a preposition.

The present section is organized as follows. I will begin by briefly presenting in Section III.4.1.2.2.1.1.1 Nagaraja’s (1985) and Rabel’s (1961) points of view on the patterns of use of the interrogatives *no* and *ey*. In Section III.4.1.2.2.1.1.2, I will discuss the use of *ey* and *no* with gender-number markers and possible correlations with the opposition human vs. non-human. In Section III.4.1.2.2.1.1.3, I will examine the use of *ey* and *no* with the elements preceding and directly relating to them other than gender-number markers and the way the

interrogatives and these elements interact with respect to the opposition human vs. non-human. Finally, in Section III.4.1.2.2.1.1.4 a brief summary will be provided.

4.1.2.2.1.1.1 Nagaraja (1985) vs. Rabel (1961)

Both major sources on Standard Khasi I consulted, Nagaraja (1985) and Rabel (1961), speak about the existence of ‘who?, what?’ interrogative pronominals in Standard Khasi. However, the two interpretations differ considerably on certain points, most importantly in the way they analyze the difference between *no* and *ey*.

Nagaraja (1985:12) claims that *no* means ‘which (one(s))?’, i.e. it is selective, or it means ‘who?’, while *ey* is just ‘who?’ or ‘what?’. Basing himself on Nagaraja’s description, Bhat (1989:479) prefers to speak about a “distinction between a definite domain (*no*) and indefinite domain (*ey*)”. There are some contexts indeed where usually *no* appears to differ from *ey* in being selective. For instance, when the two are used attributively (8a, 8b) vs. (8c) or in reported questions (9a) vs. (9b).¹⁴ However, on the whole, the differentiation between *no* and *ey* as selective vs. non-selective is best viewed as a tendency rather than a rule. (This tendency might have been a rule at some earlier stage). In non-standard varieties only reflexes of the second root are found (7).

Standard Khasi

- (8) a. *u no u khinna² u dey u ba staad?*
 M.SG IPW M.SG boy M.SG be M.SG REL intelligent
 ‘Which boy is intelligent? [The boy who came here is intelligent]’
 (Nagaraja 1985:111)
- b. *na ka no ka linti phi la leyt?*
 from F.SG IPW F.SG way 2.HON PST go
 ‘From which way did you go?’ (Nagaraja 1985:13)
- c. *na ka ey ka jynthmu phi kwa² ya ka?*
 from F.SG IPW F.SG purpose 2.HON want OBJ 3SG.F
 ‘For what purpose do you want it?’ (Nagaraja 1985:13)

¹⁴ Perhaps, also when *no* and *ey* are used as the base for indefinite pronominals. Thus, Nagaraja (1985:104) gives [*u* (or *ka/i/ki*) + *ey* + *u* (or *ka/i/ki*) + *ey*] ‘whatever, something, anything’ vs. [*u* (or *ka/i/ki*) + *no* + *u* (or *ka/i/ki*) + *no*] ‘whoever, whichever, anyone’. However, Nagaraja does not provide any examples, and Rabel (1961) does not seem to provide any support for Nagaraja’s interpretation either.

- (9) a. *u kwa[?] ba n tip u no u ba la kren*
 3SG.M want REL FUT know M.SG REL M.SG REL PST speak
 ‘He wants to know which (of them) spoke’ (Nagaraja 1985:12)
- b. *u kwa[?] ba n tip u ey u ba la kren*
 3SG.M want REL FUT know M.SG REL M.SG REL PST speak
 ‘He wants to know who spoke’ (Nagaraja 1985:12)

Rabel (1961) has a different view on the differentiation between *no* and *ey*. According to Rabel (1961:67-69), *no* and *ey* “both mean ‘who?, which?, what?’” and “seem to be in free variation”, although “in practice /-²ey/ is used more often with personal pronouns [i.e., with the gender-number markers: *u* M.SG, *ka* F.SG, *i* DIM and *ki* PL] and prepositions”. Note, however, that when Rabel speaks about the preference for *ey* with “prepositions”, she explicitly restricts herself to the “directive prepositions” /na/ ‘from’, /ha, ša/ ‘to’ and the “preposition” /kum/ ‘like’, as illustrated in (10). I will further elucidate on the use of “prepositions” and various preposed elements, other than gender-number markers, with *ey* and *no* in Section III.4.1.2.2.1.1.3.

Standard Khasi

- (10) a. *na[?] ey* (or *nu*)?
 from IPW
 ‘from where?, from whom?’ (Rabel 1961:69)
- b. *ša[?] ey* (or *nu*)?
 to IPW
 ‘to where?, to whom?’ (Rabel 1961:69)
- c. *da Ei* (or *no*)?
 with/by IPW
 ‘with what?, by whom?’ (Subbarao & Temsen 2003:200)
- d. *kum ey* (or *nu*)?
 like(a manner) IPW
 ‘how?’ (Rabel 1961:69)¹⁵

4.1.2.2.1.1.2 Gender-number markers and the distinction human vs. non-human

Let us examine the meanings that *no* and *ey* may have when marked for gender-number with preposed *u* M.SG, *ka* F.SG, *i* DIM or *ki* PL. Recall that in fact *no* is rarely used with gender-number markers (Rabel 1961:69; Section III.4.1.2.2.1.1.1), as compared to its frequent use with the other preposed elements (cf. Section III.4.1.2.2.1.1.3). When *no* is marked for gender-number, it

¹⁵ Nagarja (1985:47) and Subbarao & Temsen (2003:200) give only the form *kum-no*.

appears to differ from *ey* in being preferably selective as in (8a), (8b) and (9a) above, although not necessarily when it is used as an interrogative pronominally. Consider, for instance, (11-13). It should be mentioned, however, that at least in (11) and (12) Nagaraja himself translates *ka no* as ‘which?’, i.e. ‘Which is that?’ (over there) and ‘Which is your name?’. Yet, this is not really a normal way of asking these questions in English.

Standard Khasi

- (11) *ka ey/no ka tay?*
 F.SG IPW F.SG DEM
 ‘What is that?’ (over there) (Nagaraja 1985:12)
- (12) *ka ey/no ka kirteŋ joŋ phi?*
 F.SG IPW F.SG name of 2.HON
 ‘What is your name?’ (Nagaraja 1985:13)
- (13) *u ey/no u tu u briew?*
 M.SG IPW M.SG DEM M.SG man
 ‘Who is that man?’ (Nagaraja 1985:12-13)

As such, neither *ey* nor *no* is restricted to the human or non-human interpretation. Therefore, it seems reasonable to concentrate on the semantics of the gender-number markers that *ey* and *no* take. Recall that on the whole, the Standard Khasi gender-number system is indifferent to the opposition the human vs. non-human (Section III.4.1.2.1), which in the case of the interrogative pronominal *ey* gives us the following result: *u ey* ‘who?, what? (M.SG)’, *ka ey* ‘who?, what? (F.SG)’, *i ey* ‘who?, what? (DIM.SG)’ and *ki ey* ‘who?, what? (PL.SG)’, *i ey* ‘who?, what? (DIM.PL)’ (Rabel 1961:68). Still, in practice certain correlations between the gender-number of the interrogative and the meanings ‘who?’ and ‘what?’ do exist. Thus, the meaning ‘what?’ appears to be most often expressed by the feminine singular form *ka ey* (or *ka no*), as in (11, 12, 14, 20).¹⁶

Standard Khasi (Nagaraja 1985:115, Subbarao & Temsen 2003:209)

- (14) *(ya) ka ey phii kwa??*
 OBJ F.SG IPW 2.HON want
 ‘What do you want?’

This use is clearly due to the fact that in Standard Khasi non-human controllers of which the gender is unknown trigger the feminine agreement pattern, the feminine showing up also when an atypical controller is used or no controller is

¹⁶ The phrase “most often” here should be understood in respect to the gender-number marked interrogatives, because there is also another ‘what?’ interrogative *ayu* (see Section III.4.1.2.2.1.2).

possible at all (cf. Section III.4.1.2.1).¹⁷ In contrast, most often the meaning ‘who?’ appears to be expressed by the masculine singular form *u ey* (or *u no*) (13, 15),¹⁸ at least in situations where the speaker does not assume that the person inquired about is a woman. In the latter case, it appears that the feminine form *ka ey* (or *ka no*) must be used, just as with non-humans.

Standard Khasi (Nagaraja 1985:12)

- (15) *u ey u ba nuud ba n wan ša ne?*
 M.SG IPW M.SG REL dare REL FUT come to DEM
 ‘Who dares to come here?’

That the masculine agreement pattern is the default option for questions about persons may be further corroborated by the fact that apparently only the masculine singular marker *u* may be omitted in (16a), but not the feminine singular *ka* in (16b) or the plural *ki* (16c). According to Rabel (1961:128-129), the gender-number marker before the relative *ba* may be left out “at the discretion of the speaker”, but apparently in questions about persons this is only possible when the speaker presupposes that the person is a man (16a) or when the speaker does not have any particular presupposition on the gender-number of the person in issue (17).

Standard Khasi (Subbarao & Temsen 2003:204-205)

- (16) a. *ma no (u) ba wan?*
 EMPH IPW M.SG REL come
 ‘Who came? (the speaker presupposes that the person who came is a man)’
 b. *ma no ka ba thO?*
 EMPH IPW F.SG REL come
 ‘Who wrote? (the speaker presupposes that the person who wrote is a woman)’

¹⁷ Rabel (1961:94) also notes that the feminine “is used for neuter gender in reference to an unknown person or thing: /ka ʔey/ ‘who?, what?’”. This should probably be interpreted in the sense that when the speaker has no idea whether the referent is human or non-human, non-human would be the default option and consequently the speaker would use the feminine form of the interrogative, since it is the default option with non-humans. In this respect, Standard Khasi is in fact not that different from English, which seems to prefer *what?* in such situations.

¹⁸ Given that ‘who?’ is also often expressed by means of a combination [the emphatic marker *ma* + the interrogative pronominal *no*] (see (16, 17), as well as examples (22c, 23) and the discussion thereafter), “most often” here means most often as regards the gender-number marked interrogatives

c. *ma no ki ba thya?*

EMPH IPW F.SG REL sleep

‘Who slept? (the speaker presupposes that there are several persons who slept)’

(17) *ma no ba lEit ša yOw?*

EMPH IPW REL go to market

‘Who went to the market?’

As to the diminutive(-honorific) form *i ey*, it can mean ‘who?’ or ‘what?’ (singular or plural) depending on the context (18, 19), because the diminutive does not correlate with (non-)humanness in the same way as the masculine and the feminine do.

Standard Khasi (Anvita Abbi 2003, see <http://listserv.linguistlist.org/cgi-bin/wa?A2=ind0303b&L=lingtyp&D=1&P=326>)

(18) *ii ey bə hap?*

DIM IPW REL fall

‘Who/what fell? (diminutive or honorific; used for needle, baby, mouse, mice, etc.)’¹⁹

(19) *ii ey i wən?*

DIM IPW DIM come

‘Who came? (honorific, i.e. parents)’

Given that in the plural there is no distinction between the masculine and the feminine, the plural interrogative *ki ey* also means ‘who?’ or ‘what?’ depending on the context.

Summing up, (i) the masculine singular form is typically used as ‘who?’, while (ii) the other three gender-number forms are used as both ‘who?’ and ‘what?’ (even though the meaning ‘what?’ appears to be most often expressed by the feminine singular form).

4.1.2.2.1.1.3 “Prepositions” and the distinction human vs. non-human

As has been mentioned in Section III.4.1.2.2.1.1.1, for Rabel (1961) prepositions are only the “directive prepositions” /na/ ‘from’, /ha, ša/ ‘to’ and the “preposition” /kum/ ‘like’. I will use the term “preposition” in a broader sense to refer to both to Rabel’s prepositions and what she calls “other words and bases”

¹⁹ Abbi herself gives the verb as *bəhap’*. The gloss has been added by me. Note that the final apostrophe in Abbi’s *bəhap’* may be a typo. In the sources I consulted the only similar looking verb meaning ‘fall’ I have encountered was Pnar Khasi *haap* (Grierson 1906:II:29). Besides, it is difficult to see what this apostrophe could stand for phonologically in Khasi.

(1961:68-69), i.e. to any element preceding and directly relating to the interrogatives *no* and *ey* other than a gender-number marker. Note in this respect that most of these other “words and bases” can actually be safely analyzed as prepositions as well.²⁰

When *no* and *ey* are preceded by a preposition, the gender-number marker is regularly omitted. There appear to be only two exceptions. First, the gender-number marker is preserved when *no* and *ey* are used as preposed attributes. In this case, the preposition heads the whole NP (8b, 8c), i.e. it does not relate directly to the interrogative. Second, the gender-number is preserved with *ey* (but not with *no*), when the preposition is *ya*, which is sometimes used to mark the direct object or the beneficiary/recipient (for more details see below), as in (20).

Standard Khasi (Subbarao & Temsen 2003:201)

- (20) *u jon u pule ya ka Ei*
 M.SG John 3SG.M read OBJ F.SG IPW
 ‘What did John read?’

According to the data in Rabel (1961), it appears to be possible to classify the elements preceding and directly relating to the interrogatives *no* and *ey*, other than gender-number markers, in two groups: (i) those that favour *ey*, which are summarized in (21), and (ii) those that favour *no*, which are summarized in (22).

- (21) *The “prepositions” with which ey is preferred to no in Standard Khasi (based on Rabel 1961:68-69)*

- a. the “preposition” /na/ ‘from’ (10a)
- b. the “prepositions” /ša, ha/ ‘to’ (10b)
- c. the “preposition” /kum/ ‘like’ (10d)

- (22) *The “prepositions” with which no is preferred to ey in Standard Khasi (primarily based on Rabel 1961:68-69)²¹*

- a. the “locative bases” /haŋ-/ ‘in, at (a place)’ and /naŋ-/ ‘from (a place)’, as in /haŋ-nu/ ‘in what place?’ and /naŋ-nu/ ‘from what place?’
- b. the “verbal base” /kat-/ ‘(be) so much/many (as)’, as in /kat-nu/ ‘how much?, how many?’, (30)
- c. the “emphatic” /ma-/ , as in /ma-nu/ ‘who?’ (23)
- d. the close future “auxiliary” /la/ ‘at (a time in the future)’, as in /la-nu/

²⁰ Prepositions are also sometimes claimed to be case markers (e.g., Subbarao & Temsen 2003), but I do not find this analysis convincing. It should be mentioned that unlike with *no* and *ey*, with other nominals the gender-number marker is usually preserved after a preposition, as in *joŋ u briiw* ‘of a man’ (Grierson 1906:II:8).

- ‘when (in the future)?’ (24)
- e. the “temporal prefix” /mn-/ ‘at (a time in the past or present)’, as in /mn-nu/ ‘when (in the past)?’ (25)
- f. the possessive preposition *joŋ* ‘of’, as in *joŋ no* ‘of whom?, whose?’ (26)
- g. the preposition *ya*,²² sometimes marking the direct object or the beneficiary/recipient, as in *ya no* ‘who?, whom?’ (27)
- h. the preposition *bad* (also written as *bat*) ‘with (accompaniment)’ (also conjunction ‘and’), as in *bad no* ‘with whom?’ (28)

Standard Khasi

- (23) a. *dey ma nu? dey ma 'uu*
be EMPH IPW be EMPH 3SG.M
‘Who is it? It’s him’ (Rabel 1961:138)
- b. *ma no (u) ba wan?*
EMPH IPW M.SG REL come
‘Who came?’ (Subbarao & Temsen 2003:205)
- (24) *min no u la wan?*
at(a time in the past or present) IPW 3SG.M PST come
‘When did he come?’ (Nagaraja 1985:43)
- (25) *la no u n wan?*
at(a time in the future) IPW 3SG.M FUT come
‘When will he come?’ (Nagaraja 1985:43)
- (26) *ka dey ka kot joŋ no?*
3SG.F be F.SG book of IPW
‘Whose book is (it)?’ (Nagaraja 1985:57)
- (27) *phi la khot ya no?*
2.HON PST call to IPW
‘Whom did you call?’ (Nagaraja 1985:49)
- (28) *phi leyt bad no?*
2.HON go with IPW
‘With whom did you go?’ (Nagaraja 1985:55)

²¹ The last three lines (22f-h) have been added by me, because Rabel (1961:67-69) does not mention these cases explicitly.

²² Even though according to Subbarao & Temsen (2003:208-209), *ya* may be common with *ey* as well, I give it here because when introduced by *ya*, *ey* is first marked for gender-number (cf. (20) above). Consequently, the semantics of the gender-number marker is then more important for the interpretation of the interrogative *ey* than the semantics of the preposition.

The data summarized in (21) and (22) appear to support Rabel's claim that *no* and *ey* "both mean 'who?', which?, what?'". To begin with, the prepositions summarized in (21) and (22) do not seem to favour the selective or the non-selective interpretation of *no* and *ey*. This remains a matter of context. Similarly, whether *no* and *ey* get a human or a non-human interpretation in combination with the locative prepositions *na* 'from' (21a) and *ha*, *ša* 'to' (21b) seems to depend on the larger context. At the same time, the non-human interpretation of *no* and *ey* in (21c) and of *no* in (22a), (22b), (22d), and (22e) is fully determined by the semantics of the "preposition" it is used with, i.e. /kum/, /haŋ/ and /naŋ/, /kat/, /la/ and /mn/ respectively. Similarly, the emphatic marker *ma* favours the human interpretation 'who?'. The reason is that the emphatic marker *ma* is otherwise restricted to use with personal pronouns (Nagaraja 1985:102), as in *ma u* 'he (and not someone else)'.²³

The prepositions *joŋ* 'of' (22f), *ya* 'to, for, object marker' (22g) and *bad* 'with, and' (22h) also appear to favour the human interpretation of *no*. However, the explanation here is somewhat less straightforward than in the previous cases because when used with regular nominals, the three prepositions are not restricted to nominals with animate, let alone human referents. What seems to happen here is that in the context of a question, i.e. with an interrogative pronominal as the dependent, it is the most specific meaning of the preposition that is preferred. In turn, the latter meaning happens to favour the human interpretation 'who?' of the interrogative pronominal. For instance, with *joŋ* 'of' it is its possessive meaning that is favoured and consequently, the interrogative pronominal is interpreted as a possessor, which is typically a human.

As to the preposition *ya*, according to Rabel (1961:76), it "precedes nouns and pronouns which are the goals of action and which are called 'objects'". Interestingly, "in practice /ya/ is omitted as often as it is used" and "only in the case of two objects one of them is always preceded by /ya/", as in (29).

Standard Khasi

- (29) *ʔuu hiikay ya ŋa ka ktien pharaŋ*
 3SG.M teach to 1SG F.SG language English
 'He teaches me English' (Rabel 1961:77)

Of the two objects, it seems then to be typically the beneficiary/recipient that gets marked by *ya*. When no verb is present at all, *ya* also typically seems to be reserved for beneficiaries/recipients, as in (30).

²³ Subbarao & Temsen (2003:201) call *ma* "the nominative case marker". However, even if one follows this analysis, the so-called "nominative" marker will still be restricted to a use with personal pronouns (for focalization purposes).

Standard Khasi

- (30) *ya phii kat nu, Koŋ Yo?*
 for 2.HON so.many.as IPW sister PROP

‘How many (shall I prepare) for you, sister Yo?’ (Rabel 1961:127)

Compare also examples (1) and (26), where the verb is the same, *dey* ‘be’, but the preposition *ya* is used only for the beneficiary/recipient in (1), viz. *ya ŋaa* ‘for me’. In other words, the most specific function of *ya* appears to be the marking of beneficiaries/recipients, which are typically humans.

As to *bad*, it can be used as a conjunction ‘and’ and as a preposition ‘with’, in the meaning of accompaniment, not instrument. With the interrogative pronominal *no*, *bad* seems to be used as a preposition. Apparently, the most specific interpretation of the prepositional *bad* is ‘in the company of’, with the companion being normally a human.

4.1.2.2.1.1.4 The interrogatives *no* and *ey* and the distinction human vs. non-human: summary

As demonstrated above, the difference between the interrogatives *no* and *ey* does not depend directly on the opposition human vs. non-human. On the whole, it appears that *ey* is preferred in contexts where normally the interrogative must be marked for gender-number, whereas *no* is preferred in contexts where gender-number marking of the interrogative is not allowed. The exact interpretation of *no* and *ey* as ‘who?’ or ‘what?’ (or either) mostly depends on two factors, the overall context and/or the meaning of the element preceding the interrogative and directly relating to it. In the latter case, the most important distinction appears to be between the gender-number markers and the rest, where “the rest” typically implies a preposition. The usual relations between *no* and *ey* and the meanings ‘who?’ and ‘what?’ are summarized in Table 1.

4.1.2.2.1.2 The interrogative *ayu*

The interrogative pronominal *ayu* ‘what?’ “usually occurs at the end of a clause” (Nagaraja 1985:12) and it seems to be restricted to the direct object (31, 32), predicate nominal (33, 34) and attribute (35) functions. Furthermore, it cannot be marked by a preposition or a gender-number marker.

Table 1. An overview of the relations between the interrogative pronominals *ey* and *no* and the meanings ‘who?’ and ‘what?’ in Standard Khasi

Gender-number marking	Usual IPW	Usual meaning	
		Human (‘who?’)	Non-human (‘what?’)
present	M.SG <i>u</i>	✓ (man or sex unknown)	
	F.SG <i>ka</i>	✓ (woman)	✓
	DIM <i>i</i>	✓ (diminutive, honorific)	✓
	PL <i>ki</i>	✓ (plural)	✓
absent	<i>no</i>	✓	✓

Standard Khasi

- (31) *phii kwa[?] [?]ayuu?
 2.HON want what
 ‘What do you want?’ (Rabel 1961:35)*
- (32) *u la le[?] ayu?
 3SG.M PST do what
 ‘What did he do?’ (Nagaraja 1985:13)*
- (33) *u siŋ u loŋ ayu?* [*u siŋ u loŋ u doktor*]
 M.SG PROP 3SG.M be what
 ‘What is Singh? [Singh is a doctor.]’ (Nagaraja 1985:111)
- (34) *phi kirteŋ ayu?*
 2.HON name what
 ‘What is your name?’ (Nagaraja 1985:13)
- (35) *phi-n lej šuuwa ša ka basa [?]ayuu?
 2.HON-FUT go first to F.SG store what
 ‘[A and B have arrived at the market. A asks B:] To what/which store will you go first? [B: ‘We had better look at the meat first, hadn’t we?’ They go to a meat stand a look at the meat]’ (Rabel 1961:237-238)*

Usually, it appears to be possible to use *ey* or *no* instead of *ayu*, as in (36) vs. (31), (37) vs. (32), (38) vs. (34), (39) vs. (35), whereas the possibilities of replacement in the opposite direction appear to be much more restricted.

Standard Khasi

- (36) *(ya) ka ey phii kwa'?*
 OBJ F.SG IPW 2.HON want
 'What do you want?' (Nagaraja 1985:115, Subbarao & Temsen 2003:209)
- (37) *ka ey ka dey ŋa le'?*
 F.SG IPW F.SG be 1SG do
 'What shall I do?' (Nagaraja 1985:116)
- (38) *ka ey/no ka kirteŋ joŋ phi?*
 F.SG IPW F.SG name of 2.HON
 'What is your name?' (Nagaraja 1985:13)
- (39) a. *na ka no ka linti phi la leyt?*
 from F.SG IPW F.SG way 2.HON PST go
 'From which way did you go?' (Nagaraja 1985:13)
- b. *na ka ey ka jingthmu phi kwa' ya ka?*
 from F.SG IPW F.SG purpose 2.HON want OBJ 3SG.F
 'For what purpose do you want it?' (Nagaraja 1985:13)

Besides the aforementioned formal restrictions on its use, the interrogative *ayu* appears to be preferred in more abstract contexts. Thus, it is used to interrogate the kind of action, as in (32). Similarly, (31) seems to be a rather general inquiry on what the interlocutor would like to do or to happen. Example (33) inquires about the class of persons that Singh belongs to rather than his identity, as in (40).

Standard Khasi (Nagaraja 1985:111)

- (40) *u ey u dey u jon? [u jon u dey u paralok joŋ ŋa]*
 M.SG IPW 3SG.M be M.SG PROP
 'Who is John? [John is a friend of mine.]'

The use of *ayu* in (34) to inquire about someone's name also fits its overall pattern of use rather well. It is comparable, for instance, to the use of a manner interrogative *kak* 'how?' rather than *čto* 'what?' in questions about someone's name in Russian (41).

Russian (Indo-European, Slavic; Russia)

- (41) *kak/*čto ego imja?*
 how/*what his name
 'What's his name?'

Note that when used as an attribute (42), *ayu* follows the noun it modifies similarly to adjectives, but unlike attributively used interrogatives *no* and *ey* (8), demonstratives or numerals.

Standard Khasi (Rabel 1961:237-238)

- (42) *phi-n lej šuuwa ša ka basa ʔayuu?*
 2.HON-FUT go first to F.SG store what
 ‘[A and B have arrived at the market. A asks B:] To what/which store will you go first? [B: ‘We had better look at the meat first, hadn’t we?’ They go to a meat stand a look at the meat]’

Morphosyntactically, adjectives in Khasi are much closer to verbs than to nominals, so that Nagaraja (1985:26) even groups adjectives together with verbs within the larger class of verbals.

Summing up, *ayu* does not seem to be a prototypical ‘what?’ as defined in Section I.2.6. Rather, it seems to be predominantly specialized in questions about abstract things, viz. things said, thought, done, etc., and as an avoidance strategy in KIND-questions (cf. Section II.2).

4.1.2.2.2 Major non-standard Khasi varieties: War, Pnar and Lyngngam

In this section I will examine the interrogative pronominals of the three major non-standard Khasi varieties. The forms have been presented in (7) and are reproduced here as (44).

- (44) *Interrogative pronominals of the three major non-standard Khasi varieties (based on Grierson 1906:II; Nagaraja 1996)*

War	(gender-number marker) + <i>ai</i> ‘who?, what?, which?’
Pnar	gender-number marker + <i>i</i> ‘who?, what?, which?’
Lyngngam	<i>yət</i> or <i>u</i> (M.SG) + (<i>i</i>) <i>et</i> ‘who?’ <i>mət/met</i> ‘what?’

Given that all data come from one old and rather sketchy (but still quite good) source, Grierson (1906:II), the discussion will necessarily be somewhat summary. I will start with War, which seems to be most similar to Standard Khasi as far as its interrogative pronominal system is concerned. Then I will proceed to Pnar, which may actually be very much like War, but the data at my disposition are rather rudimentary. I will finish by considering Lyngngam, which appears to be quite different from both Standard Khasi and the other non-standard varieties.

4.1.2.2.2.1 War

According to Grierson (1906:II:31), War uses only the interrogative pronominal *ai* “to which the appropriate article is prefixed according to gender”. Given that the gender system of War is rather like that of Standard Khasi, my expectation

would be that the masculine form *u ai* will typically be used as ‘who?’ and the feminine *ka ai* will usually be used both as ‘who?’ (woman) and as ‘what?’. At least the latter meaning of *ka ai* is explicitly mentioned by Grierson (1906:II:45). As to the diminutive, recall that War would often use the diminutive in situations where Standard Khasi would use the feminine (cf. Section III.4.1.2.1). Therefore, in War the diminutive is probably used more frequently than in Standard Khasi as a general interrogative ‘what?’ without being restricted to small things (45-46). In (46) the use of the diminutive form *i i* is apparently due to agreement with the diminutive word *i tawiaŋ* ‘name’.

War (Austro-Asiatic, Northern Mon-Khmer, Khasian; India & Bangladesh)

- (45) *i ai i ah ni iē kat te kat te?*
 DIM IPW DIM be do 3PL.HUM so.much.as DEM so.much.as DEM
 ‘What are they doing (there so excessively)?’ (Grierson 1906:II:34)

- (46) *i ai i tawiaŋ m?*
 DIM IPW DIM name 2SG
 ‘What is your name?’ (Grierson 1906:II:55)

I also found two examples of the plural interrogative *ki ai*, (47) and (48).

War (Grierson 1906:II:57)

- (47) *u himbo ki ai u le abeh di trai m?*
 M.SG child PL IPW 3SG.M come (behind) 2SG
 ‘Whose boy comes behind you?’
- (48) *ti ki ai kti hi ei iē ile?*
 from PL IPW (?buy) (?EMPH/2PL) OBJ 3SG.NON<HUM> (?also)
ti u a dui dukan ti šnoŋ
 from M.SG REL own shop from village
 ‘From whom did you buy that? From a shopkeeper of the village.’

These examples are remarkable because normally a singular form would be expected here. For instance, a singular form is found in the same sentences in Pnar (Grierson 1906:II:56). A possible explanation could be that the plural form *ki ai* is used in (47) and (48) to avoid the necessity of choosing between the masculine and the feminine forms, *u ai* and *ka ai*, in the same way as English would use the “evasive” *they* in *When a person eats too much, they get fat* instead of the generic *he* (Corbett 1991:222).

Summing up, the interrogative pronominal system of War appears to be similar to that of Standard Khasi, the most important difference being the absence of the interrogatives *no*²⁴ and *ayu* and the tendency to use the diminutive

²⁴ This root may have been marginally preserved in the War interrogative *kat-ñiah* ‘how

form of the interrogative as a general interrogative ‘what?’.

4.1.2.2.2 Pnar

According to Grierson (1906:II:25), in Pnar the interrogative pronominal root is *i* and its masculine form *u i* is used as ‘who?’ (49), while its diminutive form *i i* is used as ‘what?’ (50).

Pnar (Austro-Asiatic, Northern Mon-Khmer, Khasian; India)

- (49) *u joŋ u i u khinnah u wa bud nadiin mii?*
 3SG.M of M.SG IPW M.SG boy M.SGREL come behind 2SG
 ‘Whose boy comes behind you?’ (Grierson 1906:II:56)

- (50) *i i pirtuit mi?*
 DIM IPW name 2SG
 ‘What is your name?’ (Grierson 1906:II:54)²⁵

However, in the wordlist on page 44 ‘what?’ is given as *ka i*, which is a feminine form. Furthermore, consider example (51) where ‘from whom?’ is written by Grierson as *nei-ī* (i.e., *nei-ii*).

Pnar (Grierson 1906:II:56)

- (51) *nei-ii thied phi ka tu?*
 from.?(DIM).IPW buy 2SG F.SG DEM
na u ba dai dukaan na šnoŋ
 from M.SG REL own shop from village
 ‘From whom did you buy that? From a shopkeeper of the village.’

As follows from the answer to this question, the preposition here must be *na* ‘from’, which seems to suggest that *nei-ī* ‘from whom?’ results from the fusion of *na* with the diminutive form of the interrogative *i i*.

All in all, the interrogative pronominal system of Pnar appears to be similar to that of Standard Khasi and particularly to that of War. The most important difference with Standard Khasi is the absence of the interrogatives *no*²⁶ and *ayu* and the tendency to use the diminutive form of the interrogative as a general

many/much?’ (Grierson 1906:II:55), where *kat* is the same root as the Standard Khasi *kat* ‘(be) so many/much as’ (22b).

²⁵ Note that the word *pirtuit* ‘name’ may be diminutive, just like *i tiwiaŋ* in War (46), so that the use of the diminutive *i i* here may be due to agreement with ‘name’.

²⁶ This root may have been marginally preserved in the War interrogatives *kat-nu* ~ *kat-won* ‘how many/much?’ (Grierson 1906:II:54), where *kat* is the same root as the Standard Khasi *kat* ‘(be) so many/much as’ (22b).

interrogative ‘what?’. The difference between Pnar and War probably lies in certain details of the patterns of use of the gender-number markers on the interrogative pronominal root.

4.1.2.2.2.3 Lyngngam

According to Grierson’s grammatical sketch of Lyngngam (1906:II:18), Lyngngam uses *u* (M.SG) + (*i*)*et* as ‘who?’ and *met* as ‘what?’ (52). However, in the “standard list of words and sentences” (1906:II:56), I have found three additional interrogative pronominal forms, *joŋ iak* ‘of whom?, whose?’ (53), *am net* ‘from whom?’ (54) and *at* ‘what?’ (55).

Lyngngam (Austro-Asiatic, Northern Mon-Khmer, Khasian; India)

- (52) *phiiaw am raw met?*
 2PL (2of) do IPW
 ‘What are you (PL) doing?’ (Grierson 1906:II:21)
- (53) *u khoondiinj joŋ iak wan ha bandon am phiiaw?*
 M.SG child of IPW come at behind of 2PL
 ‘Whose boy comes behind you?’ (Grierson 1906:II:56)
- (54) *am net phiiaw thoh ukidu?*
 from IPW 2PL buy (?M.SG.DEM)
 ‘From whom did you buy that?’ (Grierson 1906:II:56)
- (55) *at iat s mi?*
 IPW (?DIM).name to 2SG
 ‘What is your name?’ (Grierson 1906:II:54)

The interrogative *net* also occurs in Lyngngam within the interrogative *kat net* ‘how many/much?’ (Grierson 1906:II:54), where *kat* is the same root as the Standard Khasi *kat* ‘(be) so many/much as’ (13b). The use of *net* in *kat net* ‘how many/much?’ as compared to its use in *am net* ‘from whom?’ in (54) may suggest that *net* in Lyngngam should be considered as a ‘who?, what?’ interrogative.

Another source on Lyngngam that I was able to consult, Nagaraja (1996), gives somewhat different forms and suggests that there is a clear distinction between ‘who?’ and ‘what?’ in Lyngngam. Nagaraja (1996:40) gives ‘who?’ as *bəyət* and ‘what?’ as *ormat*, but further on page 46 the same interrogatives are given already as *yət* and *mət* respectively. All the examples provided on the same page use the latter forms, (56-59).

Lyngngam (Nagaraja 1996:46)

- (56) *a: kudu rəu met di?*
 Q 3PL do IPW PST
 ‘What did they do?’
- (57) *a: jutu la:r s-mət?*
 Q 3SG.M come OBJ-IPW
 ‘Why (lit.: ‘for what’) did he come?’
- (58) *a mi oŋ s-yət?*
 Q 2SG.M call OBJ-IPW
 ‘Whom did you call?’
- (59) *a: gni əkot əmbi yət?*
 Q DEM book of IPW
 ‘Whose book is it?’

Interestingly, unlike Grierson (1906:II), Nagaraja gives the human interrogative ‘who?’ without the preceding masculine marker *u*. However, this may simply be due to the fact that in Nagaraja’s examples the respective interrogative is preceded by a preposition (or prefixed case marker), i.e. it is in a context where the gender-number marking is often omitted on the interrogative pronominals in the other Khasi varieties.

The interrogatives of Lyngngam look rather different from the interrogatives of other Khasi varieties. It may be speculated that the initial *n-* of *net* is cognate to the Standard Khasi interrogative root *no*, while the *m(e)-* part of *met* ‘what?’ comes from Garo *mai-* ‘what?’, a neighbouring Tibeto-Burman language (Burling 1961:41) (60).²⁷

Standard Garo (Sino-Tibetan, Tibeto-Burman, Jingpho-Konyak-Bodo; India)

- (60) *i-a mai-a?*
 this-NOM what-NOM
 ‘What is this?’ (Burling 1961:41)

²⁷ In fact, the interrogative *met* as a whole may be of a Tibeto-Burman origin, as well as *-et* and *at* interrogatives. Thus, Standard Garo also has *badi-* ‘which [N]?’, and a few other interrogatives based on *ba(di)-* (Burling 1961:41-42). A Bangladesh variety of Garo, sometimes referred to as Mande or Mandi, bases most of its non-human interrogatives on *ba(t)-/ba(d)i-* (Burling 2004:48-49; used on its own *badi* means ‘how?, which way?’). In Rabha, which is closely related to Garo and is spoken right to the north of the latter, the interrogative ‘what?’ is *at-à* (Burling forthcoming).

4.1.2.3 On the etymology of the interrogatives *no* and *ey*

It has been demonstrated above that as such the interrogative pronominal roots *no* and *ey* do not differentiate between the meanings ‘who?’ and ‘what?’. I believe that most likely this lack of differentiation is due to their origin in a deictic or pronominal root, which was indifferent to this distinction. Thus, the interrogative *no* is reconstructed by Shorto for Proto Mon-Khmer as **nɔʔ* ‘which?, what?’ (2006, #92a), which is identical to **nɔʔ/*nɔh* ‘this’ (2006, #92).²⁸ In turn, the Khasi interrogative *ey* may be compared to a Proto Mon-Khmer pronominal root **[ʔ]aay* ‘other’ (Shorto 2006, #1435) or a deictic root **[ʔəy]ʔ/*[ʔəy]h/*h[əy]ʔ* ‘this, that’ (Shorto 2006, #1435a).²⁹ Originally, in both cases the deictic/pronominal root must have been modified by a postposed interrogative, in all probability **m[o]ʔ/*m[o]h* ‘what?’ (Shorto 2006, #136). Thus, the whole construction literally meant something like ‘what one?’, i.e. ‘which one? (person or thing)’.³⁰ Later, the interrogative modifier was lost and the deictic/pronominal element became interrogative itself. This is a development similar to Italian *che cosa?* ‘what thing?’ to *cosa?* ‘what?’ or Wa *pui mɔʔ* ‘what/which person?’ to *pui* ‘who?’ (Northern Mon-Khmer, Western Palaungic; see Section III.4.1.3). An even better parallel is provided by the Stieng interrogative *ə:n/(ʔ)ɛ:n* ‘what?, (who?), what [N]?, which [N]?’ (Eastern Mon-

²⁸ The Khasi form can, for example, be compared to Vietnamese *nào* ‘which [N]?, what [N]?’ (Mon-Khmer, Viet-Muong; Vietnam; Nguyễn 1997:30), Nyahkur *nəw* with the same meaning (Mon-Khmer, Monic; Thailand; Huffman 1971), Kontoi Blang bound root *-nɔʔʔ*, as in *kənɔʔʔ* ‘what?’, *anɔʔʔ* ‘who?’, *naŋnɔʔʔ* ‘where?’ (Northern Mon-Khmer, Western Palaungic; China & Myanmar; Paulsen 1992:178, 181). Particularly interesting is Bru *nəw/nəw*, which is both an interrogative pronominal ‘who?’ and a third person singular pronoun (Eastern Mon-Khmer, West Katuic; Laos & Vietnam; Huffman 1971; Shorto 2006, #92).

²⁹ A parallel to the Khasi interrogative is provided by the second part of Khmer *(s)ʔryʔvry* ‘what?’ (Eastern Mon-Khmer; Cambodia; Shorto 2006, #1435a), Bo Luang Lawa *ʔaʔəih* ‘what?’ (Northern Mon-Khmer, Western Palaungic; China; Shorto 2006, #1435a), Umphai Lawa *ʔuʔeh* ‘what?’ (Northern Mon-Khmer, Western Palaungic; China; Shorto 2006, #1435a). Note that Shorto believes that the second element in each of the three interrogatives above is a reflex of **[ʔəy]ʔ/*[ʔəy]h/*h[əy]ʔ* ‘this, that’ (2006, #1435a). Consider also the first part of the interrogative *ʔay-ləw* ‘who?’ or ‘who?, what?’ found in some Aslian languages (see Section III.4.1.5). A similar looking Vietnamese *ai* ‘who?’ (and similar forms in several other Viet-Muong languages) may either be of a similar origin or represent a reflex of the Proto Mon-Khmer **ʔiiʔ* ‘person’ (Shorto 2006, #2). Note, by the way, that Shorto compares **ʔiiʔ* to a Proto Austronesian third person singular pronominal **ia*.

³⁰ In this respect, consider for instance the Shwe Palaung interrogative pronominal *i-mɔ*, discussed in Section III.4.1.3, or interrogative pronominals of several Aslian languages discussed in Section III.4.1.5.

Khmer, South Bahnaric; Vietnam & Cambodia; Huffman 1971; Miller 1976; Shorto 2006, #1115; see Section III.4.1.4.2), which Shorto (2006, #1115) brings back to the third person singular pronoun *[ʔ]anʔ/*ʔən[ʔ].

4.1.3 *Palaungic languages*

According to the Ethnologue, Palaungic languages belong to the Northern Mon-Khmer branch of the Austro-Asiatic phylum. There appear to be at least two Eastern Palaungic languages with a ‘who?, what?’ interrogative, Danau and Shwe Palaung. Besides these two languages and several other varieties of Palaung, the Eastern Palaungic group includes Riang and Yinchia languages. Small pockets of Palaungic speaking communities are dispersed in the neighbouring regions of Thailand, Myanmar, China and Laos, see Map 2. Danau is a small language spoken in the central part of Myanmar, just to the west of the Inle Lake. Shwe Palaung, marked as Ta-ang on Map 2 is a relatively big language spoken mostly in Myanmar and by a small group in China.

Palaungic languages are rather poorly described. The only published source dedicated to Danau I know of is Luce’s (1965) comparative wordlist, which does not contain interrogative pronominals. My information on the interrogative pronominals of Danau comes from Grierson’s (1906:I.2:222, 224) comparative vocabulary in his *Linguistic Survey of India*. Both ‘who?’ and ‘what?’ are given as *mi* in this source (1906:I.2:222, 224).

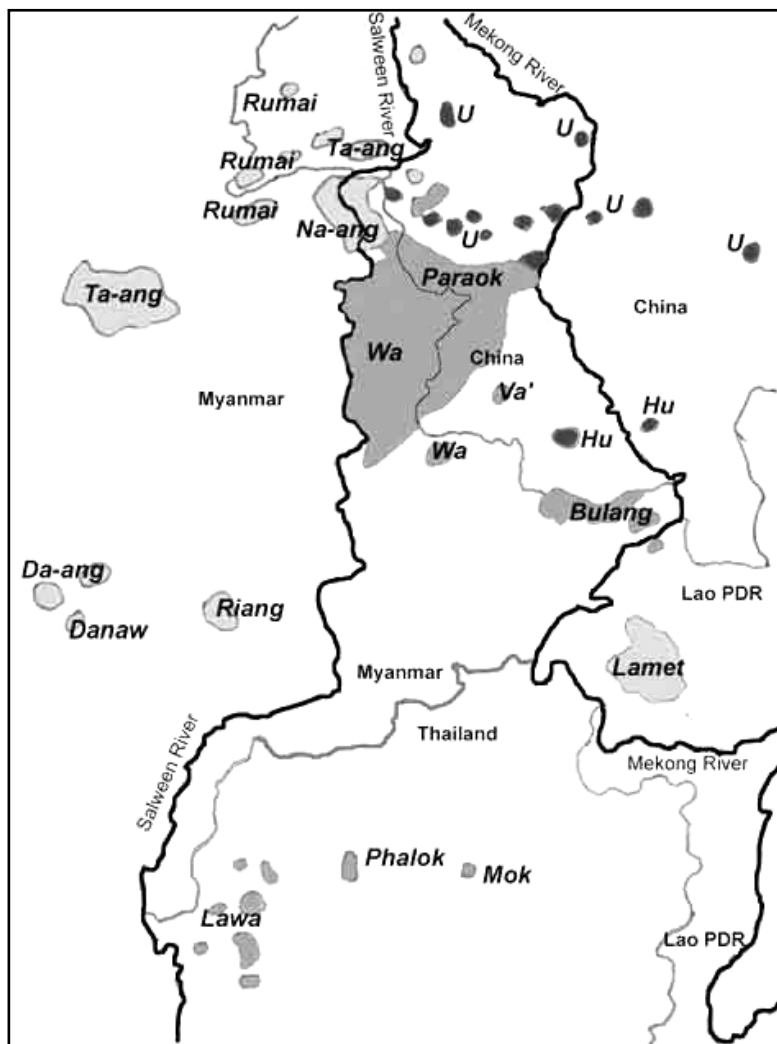
For Shwe Palaung I consulted Milne’s (1921) grammar. Milne (1921) does not explicitly mention a ‘who?, what?’ interrogative herself. However, the examples that she gives for the interrogative *im̄*³¹ suggest that it can be used not only as ‘what?’ (61) and ‘which one? (person or thing)?’ (62), but also as ‘who?’ (63), at least predicatively.

Shwe Palaung (Austro-Asiatic, Northern Mon-Khmer, Eastern Palaungic; Myanmar)

- (61) *im̄ mi hnyōm?*
 IPW 2SG believe
 ‘What do you believe?’ (Milne 1921:27)

³¹ The notation of Milne’s Shwe Palaung interrogative pronominals has been slightly modernized for citation purposes in accordance with the data in Shorto (1963:55, 2006, #46), but the original notation is preserved in the examples.

Map 2. Distribution of Palaungic languages (from <http://www.anu.edu.au/~u9907217/languages/Palaungic.html>)



Shwe Palaung (Milne 1921:28)

- (62) a. *im̄ mōh gāng mī?*
 IPW be house 2SG
 b. *im̄ gāng mī mōh?*
 IPW house 2SG be
 ‘Which (one) is your house?’
- (63) *im̄ ǎn mōh?*
 IPW 3SG be
 ‘Which is he?’, or better ‘Who is he?’

Note that although the third person pronominal *ǎn* as such is not restricted to human reference, in (63) Milne glosses it as ‘he’ and translates the examples as ‘Which is he?’. A less literal translation would, I suppose, be ‘Who is he?’. This interpretation is further supported by the fact that there are several other Northern

Mon-Khmer languages that have a cognate interrogative *mɔʔ* (and the like) meaning both ‘what [N]?, which [N]? (person or thing)’ and ‘who?’ (see below). It appears that the ‘who?, what?’ interrogative in Danau and Shwe Palaung has developed out of an interrogative literally meaning ‘what one? (person or thing)’, i.e. ‘which one? (person or thing)’. This interrogative has been based on a combination of an interrogative ‘what?’ preceded by a root meaning ‘one’ in Danau and by a demonstrative pronominal root in Shwe Palaung.

By way of comparison, consider in (64) the interrogative pronominals of Danau, Shwe Palaung and some other Palaungic languages, one Eastern Palaungic, Pale Palaung,³² and two Western Palaungic, Samtao and Wa.

(64) Danau	<i>mi</i>	‘who?, what?’
(Grierson 1906:I.2:222, 224)		
Shwe Palaung	<i>a-s^(h)e, a-s^(h)i</i>	‘who?’
(Milne 1921:19, 25-28;	<i>s^(h)e, s^(h)i</i>	‘what?’
Shorto 1963:55, 2006,	<i>mɔ</i>	‘what [N]?, which [N]? (person or thing)’, ‘what? (only with some prepositions?)’
#46)		
	<i>i-mɔ</i>	‘what?, who?, which one? (person or thing; lit. ≈ ‘what one?’)’
Pale Palaung	<i>ma-se</i>	‘what?’
(Janzen 1976:678)	<i>a-sei</i>	‘who?’
Wa	<i>pui (mɔʔ)</i>	‘who?’ (lit.: ‘person (what/which?)’)
(http://wadict.soas.ac.uk)	<i>mɔʔ</i>	‘who?’, ‘what [N]?, which [N]? (person or thing)’
	<i>(pa) tiʔ</i>	‘what?’
Samtao	<i>mo¹-mɔ²</i>	‘who?’ (< *‘person what?’) ³⁴
(Paulsen 1992:181) ³³	<i>miʔ-mɔ²</i>	‘what?’
	<i>-mɔ²</i>	‘what [N]?, which [N]? (person or thing)’, ‘what?’

³² On Map 2, Pale Palaung is marked as Da-ang

³³ The superscript numbers in the examples from Samtao indicate the tone. Unfortunately, Samtao is not represented on Map 2. The Samtao live somewhere in the southern part, or just to the south of the big Wa block on Map 2.

³⁴ The element *mo¹* seems to be a reflex of Shorto’s (2006, #139) **k[n]muʔ* ‘person, human being’, similarly to Kuy *mo*: ‘person, people’ (Eastern Mon-Khmer, West Katuic; Thailand, Laos

Of particular interest here is the recurrent interrogative root *mɔ(?)mɔʔ* and the interrogatives based on it. This interrogative is reconstructed for Proto Mon-Khmer as **m[o]ʔ/*m[o]h* ‘what?’ (Shorto 2006, #136). The preposed element *i-* in Shwe Palaung *i-mɔ* is also found on demonstrative pronominals *i-ō̄* ‘this one’ (close to the speaker), *i-dīn* ‘this one’ (close to the interlocutor), *i-tāi* ‘that one’ (distant both from the speaker and the interlocutor)³⁵ and seems to be a kind of (pro)nominalizer, which can be conveniently glossed in English as ‘one’. In all probability, it is itself of a deictic origin.³⁶ The preposed element *miʔ-* in Samtao *miʔ-mɔʔ* is in all probability a reflex of a Proto Mon-Khmer root **muuy/*muəy* ‘one’ (Shorto 2006, #1495). The Danau interrogative *mi* may then go back to a form like Samtao *miʔ-mɔʔ*, in the same way as Wa *pui* ‘who?’ goes back to *pui mɔʔ* ‘what/which person?’ or Italian *cosa?* ‘what?’ to *che cosa?* ‘what thing?’.³⁷ The only difference with the Samtao interrogative is that the original form of the Danau interrogative at issue must have had the meaning ‘which one?’ (lit.: ‘what one?’) and like Shwe Palaung *i-mɔ* or English *which (one)?* it was not restricted to questions about persons or things and has remained so. In this respect, consider also the Mintil interrogative *ʔay-ləw* ‘who?, what?’ in Section III.4.1.5 (Mon-Khmer, Aslian; Malaysia).

Reasoning from the original literal meaning ‘what one?’, the semantic development of an exclusively non-human interrogative ‘what?’, such as the Samtao interrogative *miʔ-mɔʔ*, is quite natural.³⁸ Still, a few parallels to the semantic evolution in Danau and Shwe Palaung can be found in some other Northern Mon-Khmer languages. Thus, a Proto Mon-Khmer interrogative **m[o]ʔ/*m[o]h* ‘what?’ (Shorto 2006, #136) has developed into Wa *mɔʔ* ‘who?, what [N]?, which [N]? (person or thing)’, Parauk *mɔ* ‘who?, which?’ (Northern Mon-Khmer, Western Palaungic; Myanmar & China; Shorto 2006, #136), and

& Cambodia).

³⁵ Shorto (1963:55) gives the first and the last of these forms as *iʔu* ‘this’ and *itay* ‘that’ respectively.

³⁶ Consider the Standard Khasi third person singular diminutive or honorific pronoun *i* (cf. Section III.4.1.2.1), Wa *in* ‘this’ vs. *an* ‘that’ (<http://wadict.soas.ac.uk>), as well as Proto Mon-Khmer deictic **ʔ[əy]ʔ/*ʔ[əy]h/*h[əy]ʔ* and its reflexes (Shorto 2006, #1435a).

³⁷ In a similar way, Wa interrogative *tiʔ* ‘what?’ (1) may go back to another Proto Mon-Khmer root for ‘one’ **diiʔ* (Shorto 2006, #86). Shwe Palaung *se* ‘what?’ may be of the same origin as Lawa demonstrative pronominal or modifier *se* ‘that’ (Huffman 1971). In all probability, a demonstrative meaning should also be reconstructed as original for Shorto’s (2006, #46) Proto Mon-Khmer “relative/interrogative pronoun” **[ʔ]ciʔ*, of which Shorto believes *se* ‘what?’ to be a reflex.

³⁸ And in fact, it appears to be more common among the interrogatives of this structure in Mon-Khmer languages.

Kammu-Yuan *mə̌/mə̌ʔ* ‘who?, which [N]?, what [N]?’ (Northern Mon-Khmer, Khmuic; Laos; Damrong Tayanin, p.c.; Shorto 2006, #136). Note that in all these cases the reflexes of **m[o]ʔ/*m[o]h* are also used attributively as ‘which [N]?, what [N]?’, and at least in Kammu-Yuan, the selective meaning appears to be the most prominent (Damrong Tayanin, p.c.), which may explain the further evolution to ‘who?’ (cf. Section I.2.3 on a preferential link between selective interrogatives and ‘who?’ rather than ‘what?’). Another possibility would be that forms like Kammu-Yuan *mə̌/mə̌ʔ* or Wa *mɔʔ* in the meaning ‘who?’ result from a loss of a so-called “presyllable”,³⁹ i.e. that they used to be structurally similar to the Shwe Palaung interrogative *i-mɔ*.⁴⁰

4.1.4 Bahnaric languages

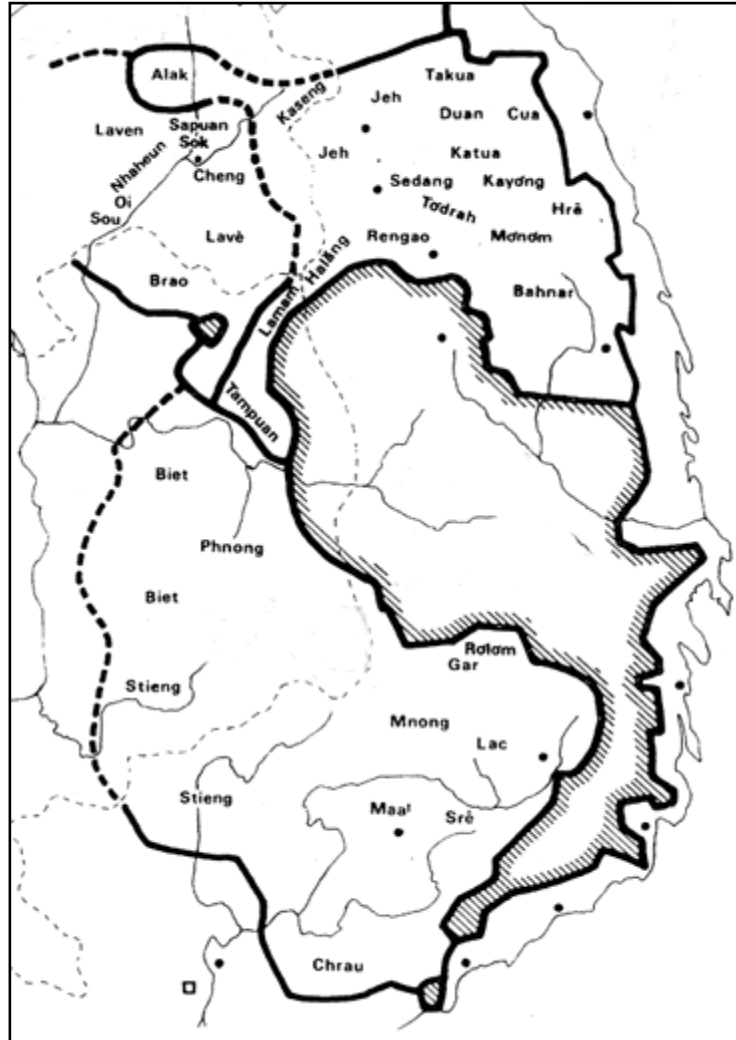
Bahnaric languages, spoken in Cambodia, Vietnam and Laos (Map 3), form one of the largest and one of the most diverse groups within the Mon-Khmer branch of the Austro-Asiatic phylum. To date, several internal classifications of Bahnaric languages have been proposed. The classifications differ both in the number of major subdivisions, ranging from three to five, and the exact affiliation of a given language. An overview of the existing classifications can be found in Sidwell (2002). Sidwell (2002) also proposes a new classification of Bahnaric languages, Figure 3. The Ethnologue divides Bahnaric into four groups, Central, North, South and West Bahnaric, which only partially correspond to those defined by Sidwell. As I am not in position to choose between the different classifications, I will arbitrarily use the Ethnologue’s labels for the purpose of reference, mentioning those suggested by Sidwell where relevant.

In the sources consulted, I found three Bahnaric languages mentioned as having an interrogative pronominal that can be used both as ‘who?’ and ‘what?’. These languages are Sapuan, Cheng and Stieng. Sapuan and Cheng are West Bahnaric both in Ethnologue’s and Sidwell’s (2002) classification, while Stieng is Central Bahnaric in Sidwell’s (2002) classification but South Bahnaric in the Ethnologue. In both cases the lack of differentiation between ‘who?’ and ‘what?’ is most likely due to the fact that the respective interrogatives have their origin in a selective interrogative pronominal ‘which one? (person or thing)?’ (literally

³⁹ The term “presyllable” in Mon-Khmer studies refers to an initial unstressed syllable of a disyllabic word which is particularly prone to various kinds of phonological erosion up to complete deletion. Presyllables are often former prefixes.

⁴⁰ The most common form of such a presyllable for ‘who?’ interrogatives (as well as some other elements, such as personal pronouns) in Mon-Khmer languages appears to be *a-*, like in Kontoi Blang *a-nɔʔ* ‘who?’ (Northern Mon-Khmer, Western Palaungic; China & Myanmar; Paulsen 1992:181) or Chut (*a-*)*ʔaɨ* (Mon-Khmer, Viet-Muong; Vietnam; Solntsev & Hoàng 2001).

Map 3. Distribution of Bahnaric languages (from Ferlus 1974 via <http://www.anu.edu.au/~u9907217/languages/Bahnaric.html>)⁴¹



something like ‘what one?’), constructed of a deictic or pronominal root (or roots), indifferent to the distinction between persons and things, and an interrogative modifier, which later disappeared. Compare similar developments discussed in Sections III.4.1.2.3 and III.4.1.3. In what follows I will first examine Sapuan and Cheng (Section III.4.1.4.1) and then Stieng (Section III.4.1.4.2) interrogative pronominals.

⁴¹ The subdivisions on this map reflect Ferlus’ own tripartite classification of Bahnaric languages. The region marked with a solid line and additional dashed downward diagonal lines confines Austronesian Chamic languages.

Figure 3. Bahnaric languages (based on Sidwell 2002; Sidwell & Jacq 2003; <http://www.anu.edu.au/~u9907217/languages/Bahnaric.html>)

West Bahnaric

- Lavi
- “the rest”:
 Jru’/Laven, Su’, Juk
 Nyaheun, Sapuan, Oi, Cheng, Laveh/Brao/Kavet/Krung

North Bahnaric

- Jeh, Halang
 - Rengao
 - Hrê, Sedang
- Unclassified: Katua, Tadrah, Manâm (Mano’m), Kayong, Duan, Kaco’, Takua

Central Bahnaric

- West Central: Kasseng, Taliang, Yaeh
- North Central: Alak
- East Central: Cua
- South Central:
 Tampuon
 Bahnar
 Chrau, Koho/Srê, Mnong, Stieng/Biet

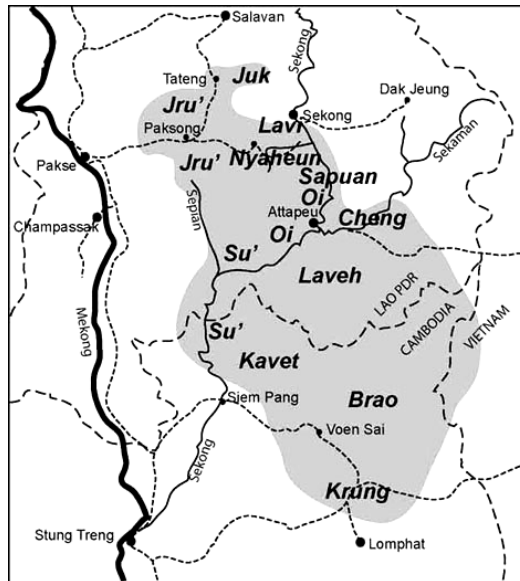
4.1.4.1 Sapuan and Cheng

Sapuan and Cheng are two rather closely related West Bahnaric languages, both in Ethnologue’s and Sidwell’s (2002) classification. Sapuan, as described by Jacq & Sidwell (1999), is spoken in the village of Ban Sapuan in the south of Laos by less than 1000 persons. Cheng (also known as Jeng or Chieng) is spoken in several villages in and around Ban Fangdeng, not so far to the south from Ban Sapuan (Sidwell & Jacq 2003:28-29). Map 4 presents the distribution of West Bahnaric languages (following Sidwell 2002).

According to Jacq & Sidwell (1999:23), Sapuan has an interrogative *ɲaj* (= *ɲaay*),⁴² “which has a broad semantic field of ‘what/which/who’ determined by the information being requested”. That is, *ɲaay* can be used pronominally as

⁴² Jacq & Sidwell (1999) leave long vowels unmarked and instead use a hacek to mark short vowels (ǎ). I preserve the original orthography of the examples, but elsewhere in this section I use *y* instead of *j* and a double vowel sign to mark vowel length.

Map 4. Distribution of West Bahnaric languages (Sidwell & Jacq 2003:3)



‘what?’ (65-66) or ‘who?’ (67-69), and attributively as ‘what [N]?, which [N]?’ (70).

Sapuan (Austro-Asiatic, Eastern Mon-Khmer, West Bahnaric; Laos)

- (65) *ʔin tənǎj səj bəm ɲaj?*
 3SG day 2SG do IPW
 ‘What are you doing today?’ (Jacq & Sidwell 1999:23)
- (66) *ʔin tənǎj səj cǎŋ ɲaj?*
 3SG day 2SG eat IPW
 ‘What did you eat today?’ (Jacq & Sidwell 1999:23)
- (67) *ɲaj dǔk ho? pəmij?*
 IPW go towards swidden
 ‘Who went to the field?’ (Jacq & Sidwell 1999:24)
- (68) *səj ma dǔk kǎp ɲaj?*
 2SG FUT go with IPW
 ‘Who will you go with?’ (Jacq & Sidwell 1999:23)
- (69) *səj kua di da ɲaj?*
 2SG sit that at IPW
 ‘Who do you live with?’ (Jacq & Sidwell 1999:24)
- (70) *kua sũk ɲaj?*
 sit village IPW
 ‘Where (lit.: ‘what/which village’) do you live?’ (Jacq & Sidwell 1999:24)

Jacq & Sidwell (1999:48) also mention that in Prachakij-karacak’s field notes

(1995 [1919]) the interrogative pronominals are given as *ŋaay* ‘what?’ and *p^huu təŋaay* ‘who?’. The latter interrogative is clearly composed of (i) *p^huu*, which must be a reflex of a Proto Mon-Khmer root **bu[u]?* ‘individual’ (Shorto 2006, #110), (ii) the preposition *tə?* ‘in, at’ (Jacq & Sidwell 1999:32), and (iii) the interrogative *ŋaay* ‘what?’. That is, the whole construction literally means something like ‘person at what?’ or ‘where person?’, i.e. ‘which person?’. In this respect, compare Nyaheun *(?)ŋeɛ* ‘where?’ vs. *klaa ŋeɛ* ‘who?’, where *klaa* is also ‘owner, master’ (Eastern Mon-Khmer, West Bahnaric; Laos; cf. <http://sealang.net>). That Prachakij-karacak (1995 [1919]) does not report the meaning ‘who?’ for *ŋaay* may be accidental, perhaps due to the way he used to collect his data. Basically, it seems to have been wordlist translation. According to Jacq & Sidwell (1999:2) Prachakij-karacak’s “data consists of a few hundred Sapuan words, seven sentences, and two paragraphs of discussion”.

The ‘who?, what?’ interrogative in Cheng has the same form as in Sapuan, i.e. *ŋaay* (Sidwell & Jacq 2003:68). Furthermore, Sidwell & Jacq (2003:68) provide an example with an interrogative of the form *pəŋaay* (71).

Cheng (Austro-Asiatic, Eastern Mon-Khmer, West Bahnaric; Laos)

(71) *sai maat pəŋaay?*
2SG name IPW

‘What is your name?’ (Sidwell & Jacq 2003:68)

In all probability, the initial *pə-* of *pəŋaay* is a reflex of the Proto West Bahnaric preposition **paa* ‘as, like’. That is, literally (71) means⁴³ ‘Your name (is) like what?’ or ‘How (is) your name?’. The use of a ‘how?’-like interrogative to question someone’s name is not uncommon cross-linguistically (cf. Sections II.3.1 and III.4.1.2.2.1.2).

Given that Cheng and Sapuan are rather closely related and that the ‘who?, what?’ interrogative in the two languages has the same form *ŋaay*, it seems reasonable to assume that in both cases the ‘who?, what?’ interrogative has the same origin. I believe that the lack of differentiation between ‘who?’ and ‘what?’ in the case of *ŋaay* is most likely due to the fact that it has its origin in a selective interrogative pronominal ‘which one? (person or thing)?’ (literally something like ‘what one?’), constructed of a deictic or pronominal root (or roots), indifferent to this distinction, and an interrogative modifier, which later disappeared. As such this development would not be unusual in Mon-Khmer. In this respect, consider for instance the Khasi interrogatives *ey* and *no* (cf. Section III.4.1.2.3 and particularly the footnotes 29-30), the Danau interrogative *mi* (Section III.4.1.3), or the Stieng interrogative *ɛ:n/(?)ɛ:n* (Section III.4.1.4.2).

⁴³ Or used to mean, depending on whether *pə* (or *pə?*, *pa?*, *paa*) still exists in Cheng as a preposition or not.

Therefore, it is necessary to demonstrate that *ŋaay* can be easily brought back to a deictic or anaphoric pronominal root. To begin with, let us consider *ŋaay*-like interrogatives in several West Bahnaric languages (72).⁴⁴

(72) *Some ŋaay-like West Bahnaric interrogatives (Jacq & Sidwell 1999:23; Sidwell & Jacq 2003:68; <http://sealang.net/mk/bahnaric.htm>)*

Sapuan	<i>ŋaay</i>	‘who?, what?, what/which [N]?’
	<i>hŋaay</i>	‘where?’ ⁴⁵
Cheng	<i>ŋaay</i>	‘who?, what?’
Laveh	<i>hŋaay</i>	‘what?’
	<i>bɪŋ ŋaay</i>	‘where?’
Oi	<i>hŋaay</i>	‘which?’
	<i>huu ŋaay</i>	‘where?’
	<i>mɛɛ ŋaay</i>	‘who?’
Su’	<i>hɬ ŋay</i>	‘where?’
Nyaheun	<i>(ʔ)ŋɛɛ</i>	‘where?’
	<i>klaa ŋɛɛ</i>	‘who?’

The initial *h-* in the forms meaning ‘where?’ must be a reflex of the Proto West Bahnaric locative preposition **hoʔ*, e.g. Su’ *hɬ* ‘at’, Sapuan *hoʔ* ‘in, at’, Brao *hɪʔ* ‘at’ (Sidwell & Jacq 2003:147-148). The initial *h-* in Laveh ‘what?’ may be a reflex of a deictic or third person singular anaphoric pronominal, such as **ʔ[əy]ʔ/*ʔ[əy]h/*h[əy]ʔ* ‘this, that’ (Shorto 2006, #1435a) or **[ʔ]anʔ/*ʔən[ʔ]* 3SG (Shorto 2006, #1435a).

As to the vowel length in the forms *ŋaay* and *ŋay*,⁴⁶ it is somewhat difficult to say whether it is the shorter or the longer forms that are closer to the original form, because in principle, both shortening of the vowel and its lengthening can be accounted for. Whatever solution is preferred, it is quite clear that *ŋ-* in *ŋa(a)y* results from a contraction of *n-k-*, where both *n-* and *k-* belong to the reflexes of deictic or anaphoric pronominal roots, such as **niʔ* ‘this’ (see Shorto 2006, #91), *nkay* ‘that’ (East Katuic; Huffman 1971), Kuy *(ŋ)kɾy* ‘there’ (West Katuic; Huffman 1971), Khasi *ka* 3SG.F (Khasian), Rengao *ga:(r)* 3SG (North Bahnaric; Shorto 2006, #26). If there was shortening of *ŋaay* to *ŋay*, then either the forms

⁴⁴ I have not included in the comparison here the interrogative *ŋaay* ‘who?’ frequent in Katuic languages (e.g., Ngeq, Dakkang, Katu, Kantu, Triw, cf. <http://sealang.net/mk/katuic.htm>) because it seems more likely to go back to Proto Mon-Khmer **[m]ŋaay* ‘person, human being’ (Shorto 2006, #1455).

⁴⁵ Note that Jacq & Sidwell (1999:24) report a different form, *da/dăw/daʔ* ‘where?’. However, this may be a loan from Vietnamese, which has *đâu* ‘where?’.

⁴⁶ The vowel in Nyaheun *(ʔ)ŋɛɛ* is clearly due to assimilation *a(a)y > ɛɛ*.

like Kuy *kɿy* should be reconstructed with a long vowel or there used to be yet another, third deictic or pronominal root following a *kV*-like deictic, e.g. **[ʔ]aay* ‘other’ (Shorto 2006, #1435) or **[ʔəy]ʔ/*[ʔəy]h/*h[əy]ʔ* ‘this, that’ (Shorto 2006, #1435a). Otherwise, the lengthening with the shift from *ɿkay* to *ɿaay* in some languages may have been compensatory.

There may be some support for the longer form as the original one outside of West Bahnaric. Consider, for instance some interrogatives of Bolyu/Lai (Mon-Khmer, Palyu; China) summarized in (73).

Bolyu/Lai (Mon-Khmer, Palyu; China; Edmondson 1995:158)

- | | | |
|------|-------------------|--------------|
| (73) | <i>ʔə-qa:iʔ</i> | ‘what?’ |
| | <i>kwəʋ-qa:iʔ</i> | ‘where?’ |
| | <i>hənʋ-qa:iʔ</i> | ‘which one?’ |

These interrogatives contain the interrogative part *qaaiʔ* with a long *a:* and apparently no traces of *n*- before *q*- to explain the lengthening. Bolyu *q*- normally seems to represent an earlier **k*-.⁴⁷

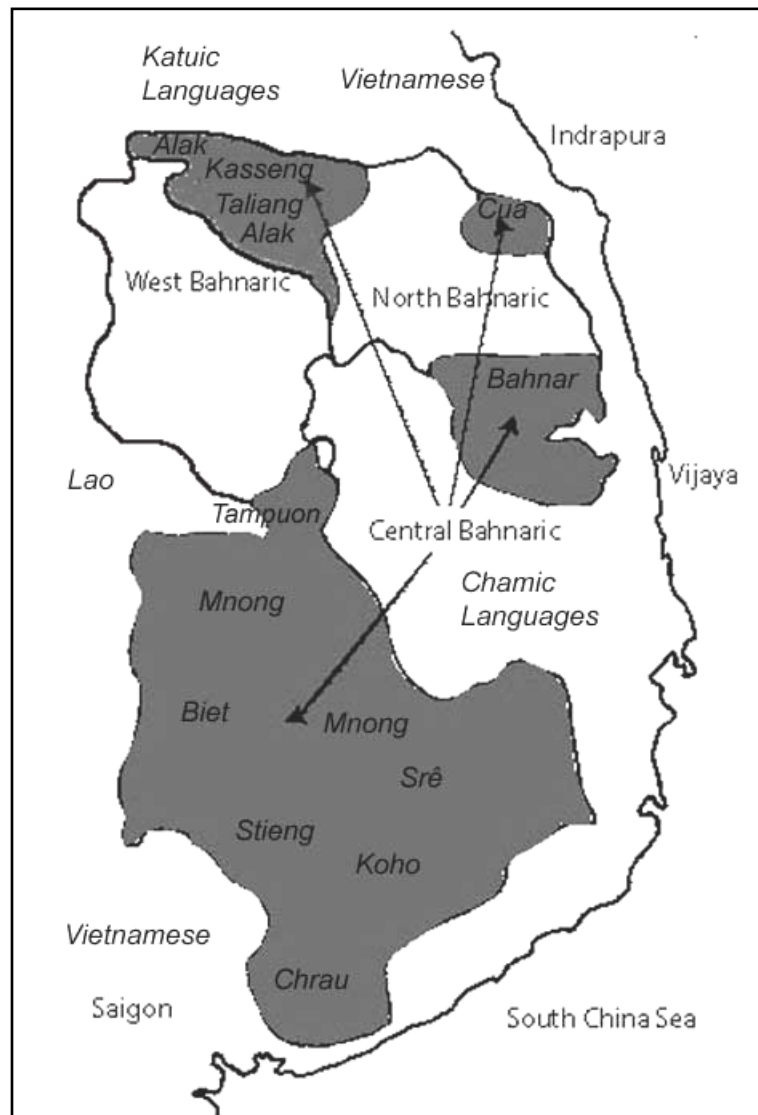
4.1.4.2 Stieng

Stieng, a Bahnaric language spoken in the south of Vietnam and in Cambodia, is classified as South Bahnaric by the Ethnologue and as South Central Bahnaric by Sidwell (2002). The Ethnologue distinguishes at least two major Stieng varieties, Bulo Stieng and Budeh Stieng, which are “different enough [...] that intelligibility is not functional”. Map 5 presents the distribution of Central Bahnaric languages in Sidwell’s (2002) classification.

I consulted three sources citing Stieng interrogative pronominals, Miller (1976), Huffman (1971) and Shorto (2006). Miller (1976) is a sketch of Bulo Stieng, as spoken in Vietnam. Huffman (1971) and Shorto (2006) are comparative wordlists, the latter also with a reconstruction attempted, which provide only forms of the interrogative pronominals without a reference to the dialect. The interrogative pronominals of Stieng cited in the three sources are summarized in (74).

⁴⁷ Consider, for instance, Bolyu *qaanʔ* ‘wings’ (Edmondson 1995:158) vs. Proto Mon-Khmer **ka(a)ŋ/*kaiŋ[ʔ]/*kiəŋ/*kaik* ‘transverse, to branch, stretch horizontally’, which has ‘wing’ as its reflex in various languages (Shorto 2006, #496). Also compare Bolyu *qatʋ* ‘to run’ (Edmondson 1995:151) with Jeh *kadaw* (North Bahnaric; Shorto 2006, #81).

Map 5. Distribution of Central Bahnaric languages
 (<http://www.anu.edu.au/~u9907217/languages/Bahnaric-Central.html>)



Stieng (Eastern Mon-Khmer, South or South Central Bahnaric; Vietnam & Cambodia)

- (74) Miller (1976:17): $\Lambda:n$ ‘what?, what [N]?’
 $pi \Lambda:n$ ‘what? (lit.: ‘what thing?’)
 $b\Lambda:n$ ‘who?’

- Huffman (1971): $? \Lambda:n$ ‘what?’
 $b? \Lambda:n$ ‘who?’

- Shorto (2006, #1115) $\varepsilon:n$ ‘what?, who?, what [N]?, which [N]?’

Unless the difference in vowels between Miller and Huffman’s forms on one hand, and Shorto’s on the other, is just a matter of notation, it may suggest that Shorto’s form comes from a different variety of Stieng, perhaps Budeh Stieng.

The initial *b-* in Bulo Stieng *b(?)a:n* ‘who?’ is clearly from the third person (singular or plural) pronoun *bu* (Miller 1976:30), which seems to be restricted to human reference and must be a reflex of Proto Mon-Khmer **bu[u]?* ‘individual, person’ (Shorto 2006, #110).

Shorto (2006, #1115) considers Stieng *a:n* ‘what?, what [N]?, which [N]?’ to reflect the same root as the third person singular pronoun **[ʔ]anʔ/*ʔən[ʔ]*. The most plausible way for it to turn into an interrogative pronoun would be to start from a phrasal selective interrogative ‘which one?’ (literally ‘what one?’) through the reduction of a postposed interrogative modifier. On the one hand, given that the pronominal at issue, **[ʔ]anʔ/*ʔən[ʔ]*, does not differentiate between human and non-human reference, it is not surprising that the interrogative based on it does not distinguish between ‘who?’ and ‘what?’ either. On the other hand, however, reasoning from the original literal meaning ‘what one?’, the semantic development of an exclusively non-human interrogative ‘what?’, as it happened in Bulo Stieng, is quite natural. In fact, it is perhaps even more expected that its development to a ‘who?, what?’ interrogative.

There are sufficient parallels to the proposed development among Mon-Khmer languages. Consider, for instance, the Khasi interrogatives *ey* and *no* (cf. Section III.4.1.2.3 and particularly the footnotes 29-30), the Danau interrogative *mi* (Section III.4.1.3), or the Sapuan and Cheng interrogatives *ɲaay* (Section III.4.1.4.1).

4.1.5 *Aslian languages*

Aslian languages are spoken by small, often still nomadic or semi-nomadic communities in the interior regions of the Malay Peninsula (cf. Map 6). These languages are usually divided into three major groups, North Aslian (or Jahaic), Central Aslian (or Senoic) and South Aslian (or Semelaic). Sometimes, an additional fourth group, including only one language Jah Hut, is extracted from Central Aslian. There appears to be only one Aslian language with a ‘who?, what?’ interrogative, Mintil. Mintil belongs to the Northern branch. The internal structure of North Aslian languages is presented in Figure 4. According to Benjamin (1976a:47), Mintil is spoken by a “small group of nomadic Negritos, numbering probably no more than 40 persons, ranging along the Tanum and coming out occasionally to Chegar Perah railway halt”. To the best of my knowledge, Benjamin’s (1976a) survey of the Aslian languages spoken in Malaysia is the only published source providing data on the interrogative pronominals of Mintil and its two closest sister languages, Batek Deq and Batek Nong.⁴⁸ These interrogatives are summarized in (75).

⁴⁸ Alternative spellings of Batek are Bateg and Bateq.

Map 6. Distribution of Aslian languages (Benjamin 1976a:46)

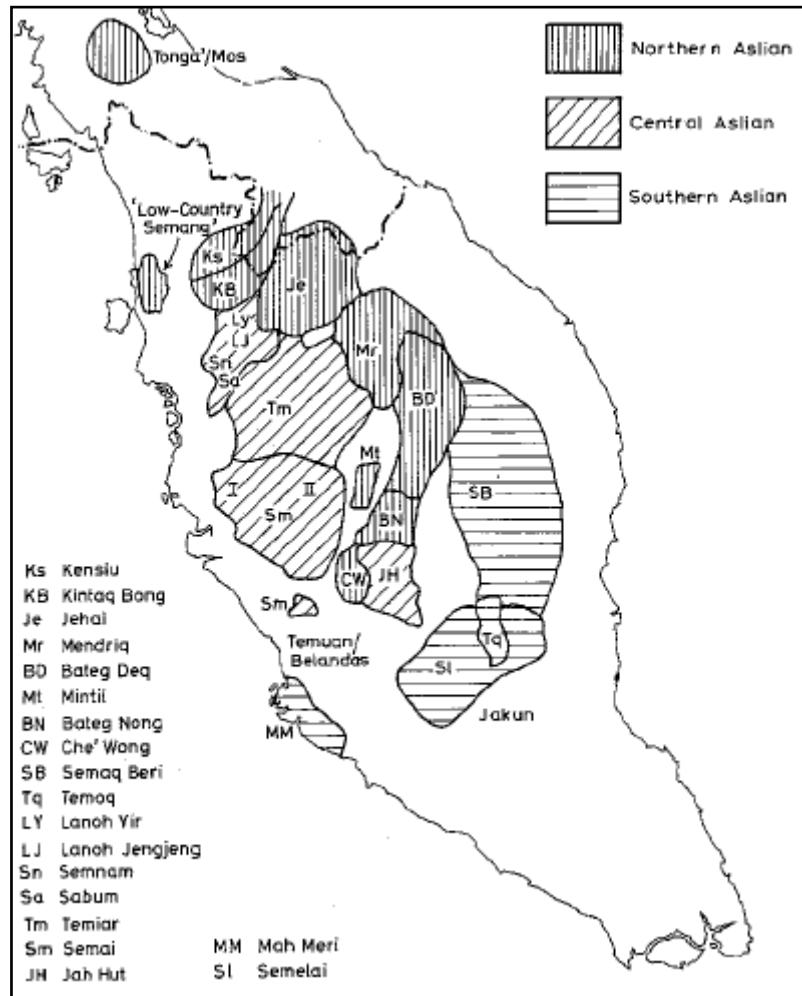
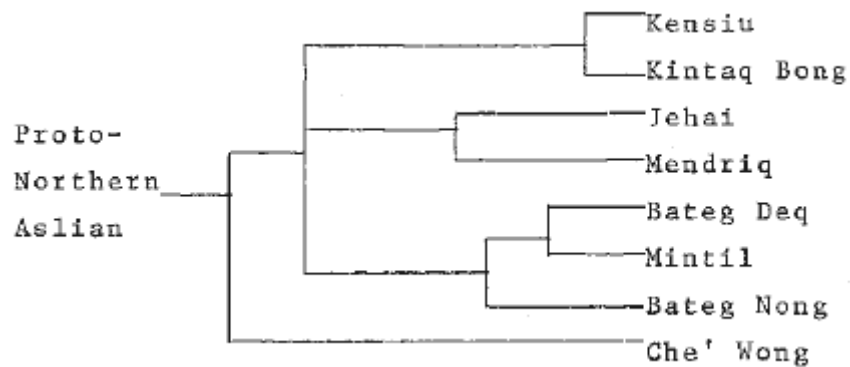


Figure 4. North Aslian languages (Benjamin 1976a:66)



(75)	Mintil	<i>?ay-ləw</i>	‘who?, what?’
	Batek Deq	<i>?ay-ləw</i>	‘what?’
		<i>?o-ləw</i>	‘who?’
	Batek Nong	<i>?ay-ləw</i>	‘what?’
		<i>tek-ləw</i>	‘who?’

The interrogative root in the forms in (75) is *ləw*. It can be compared to Kensiu *ləw* ‘what?’ (North Aslian; Benjamin 1976a:121), Temiar *lo?* ‘what?’ (Central Aslian; Benjamin 1976a:121) or Mon *lo* ‘what?, which?’ (Mon-Khmer, Monic; Myanmar & Thailand; Shorto 2006, #1885). Shorto (2006, #1885) reconstructs it for Proto Mon-Khmer as **laaw* ‘what?, which?’. The element *tek-* in Batek Nong *tek-ləw* ‘who?’ is clearly related to the word *batek* ‘person’ (Benjamin 1976a:113), i.e. originally *tek-ləw* ‘who?’ meant literally ‘what person?’. The initial *?o-* in Batek Deq *?o-ləw* ‘who?’ is a 3SG pronoun *?o?* (Benjamin 1976a:108), i.e. *?o-ləw* ‘who?’ is literally ‘what he/she?’. The human reading ‘who?’ is due to the fact that in all probability, *?o?* is restricted to human reference, as it appears to be common for third person pronominals in Aslian languages.⁴⁹

The element *?ay-* in Batek Deq, Batek Nong and Mintil interrogatives *?ay-ləw* is in all probability of deictic or pronominal origin. Consider, for instance, Proto Mon-Khmer deictic **[?əy]ʔ*?[əy]h/*h[əy]ʔ* (Shorto 2006, #1435a) and Proto Mon-Khmer **[?]aay* ‘other’ (Shorto 2006, #1435). That is, in all the three languages, Batek Deq, Batek Nong and Mintil, *?ay-ləw* must have literally meant something like ‘what one? (person or thing)’, i.e. ‘which one? (person or thing)’. Yet, while in Mintil this underspecification as to the human vs. non-human distinction has been maintained, in the two Bateks *?ay-ləw* has become restricted to a non-human meaning.⁵⁰ Remarkably, an almost complete parallel to this can be found among the Palaungic languages of the Northern Mon-Khmer branch, discussed in Section III.4.1.3. In particular, the evolution of *?ay-ləw* in Mintil can be compared to that of the Shwe Palaung interrogative *i-mɔ* and Danau *mi*, whereas the evolution of *?ay-ləw* in Batek Deq and Batek Nong is comparable to that of the Samtao interrogative *miʔ-mɔ*.

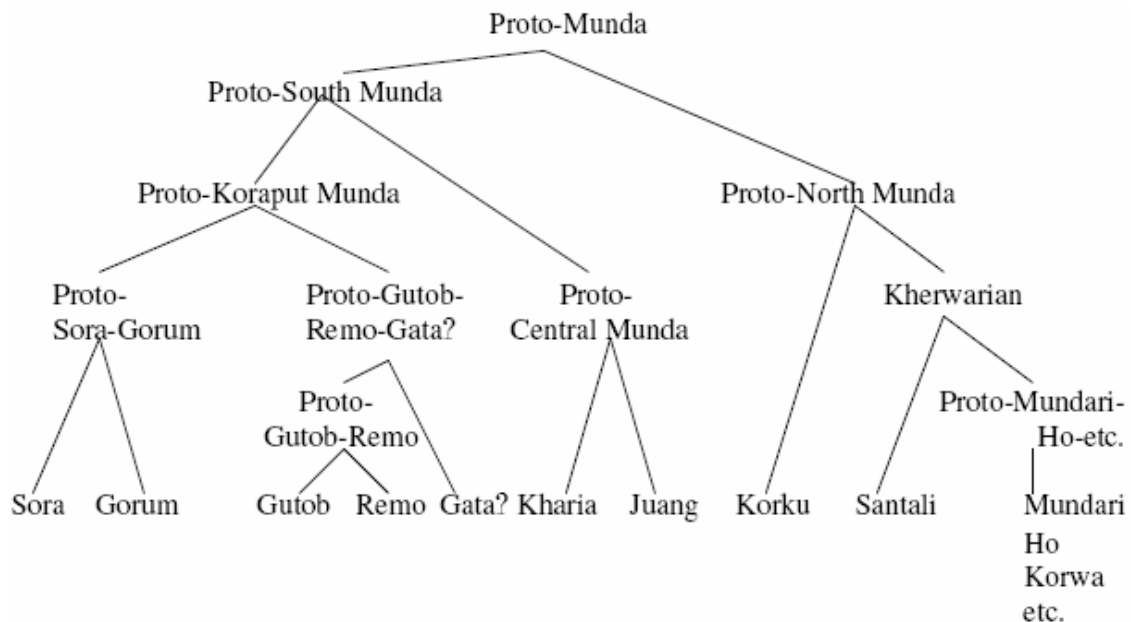
⁴⁹ In this respect, consider, for instance, Dentan (2003:6) reporting that Semai third person pronominals are restricted to human referents (Central Aslian) and Kruspe (1999:266) reporting that Semelai third person pronominals are restricted to animate referents (South Aslian).

⁵⁰ The latter has probably happened due to the similar reasons as why Batek Deq *?o-ləw* means ‘who?’ (see above).

4.1.6 Munda languages

According to the received view, the Munda languages, spoken in the centre and east of India and in Bangladesh, form the western branch of the Austro-Asiatic phylum. They are further subdivided into a southern and a northern branch. The internal structure of the two branches, as suggested by Zide (1969:412), is presented in Figure 5. On the whole, this subgrouping corresponds to the classification found in the *Ethnologue*. However, Anderson (2001) arranges the South Munda part of the tree somewhat differently. For instance, he separates Kharia from Juang and puts it closer to Gutob (also known as Bodo Gadaba) and Remo (also known as Bondo), although admitting that Kharia and Juang share many common traits.

Figure 5. *The Munda languages (Zide 1969:412)*



The relations of Juang to other South Munda languages are important for us because Juang is one of the few South Munda languages that may have a ‘who?, what?’ interrogative pronominal. The other candidates are Gorum (also known as Parenga or Parengi) and Sora (also known as Savara). According to the *Ethnologue*, Gorum may be already extinct. In the sources on North Munda languages I have encountered claims on the existence of such interrogatives only for one language, Korku.

Before proceeding further, one caveat should be noted. Given that most Munda languages, particularly of the southern branch, are rather poorly described and usually the sources available are rather sparing in illustrations of the patterns of use of the interrogative pronominals, the discussion of the interrogative

pronominal systems of these languages is necessarily somewhat preliminary.⁵¹ Of the four Munda languages mentioned, the statements citing ‘who?, what?’ interrogatives look most well-founded in the case of Juang (Dasgupta 1978; Matson 1964). In the case of Gorum, Sora, and Korku the data available is too fragmentary and not always coherent enough to allow any stronger judgements and therefore will not be discussed.⁵²

The development of ‘who?, what?’ interrogatives in Juang seems to result from the expansion of a selective interrogative pronominal ‘which one? (person or thing)’ to non-selective contexts. The latter interrogative itself is likely to go back to an interrogative meaning ‘where?’. In what follows, I will further elaborate on this issue. However, let us first examine the patterns of use of the Juang interrogatives in more detail in order to be able to decide whether we are really dealing with the ‘who?, what?’ interrogatives that we are looking for.

The interrogative pronominals of Juang are summarized in (76). The “primary meanings” are the meanings that all sources I consulted agree on. The meanings that have been mentioned only in some sources are grouped under “secondary meaning(s)” with a reference to the respective source. At the end I give interrogatives mentioned only in one source. The glosses are reproduced as found in the sources. It should be noted that the glosses ‘which?’ in (76) can be interpreted in a number of ways. First, for *aḍi/aṛi* in Juang it is not clear whether as ‘which?’ it is restricted to humans or not. Second, and more generally, it is not clear whether the gloss ‘which?’ implies the attributive selective ‘which [N]?’ , a non-selective attributive ‘what [N]?’ , or a pronominal selective ‘which one? (person or thing)’. At least in the case of Juang *maṇḍi/maṇḍi*, it is clear that a pronominal use is possible, as in (77a), where *maṇḍi* is marked with a marker *-a*, which may convey a locative, ablative or possessive meaning, or in (77b) where it is marked with *-te*, which may be a locative, ablative or instrumental marker, as

⁵¹ Of particular help for the present study have been the electronic Munda Lexical Archive compiled by Patricia J. Donegan and David Stampe (Donegan & Stampe 2004). This archive includes rather extensive lexicons, mostly of non-Kherwarian Munda languages. According to David Stampe (p.c.), the lexicons might best be considered “salvaged wordlists”, representing “(as of c. 1985) pretty much the complete lexical record”. The lexicons comprise information from most of the available published sources, as well as unpublished field notes by various researchers. I found the lexicons particularly useful. The only thing about them I deplore is the fact that both due to their format and the nature of many of the primary sources used they contain rather few examples.

⁵² The sources consulted for Gorum include Aze & Aze (1973), Donegan & Stampe (2004), Grierson (1906:IV), and Trail (1973:123, 127). For Sora, the sources are Donegan & Stampe (2004), Ramamurti (1931, 1938), [no author] (1927), and for Korku only Donegan & Stampe (2004).

(76) *Juang interrogative pronominals (based on Dasgupta 1978; Donegan & Stampe 2004; Grierson 1906:IV; Matson 1964; Mohaptra 1991)*

	Primary meaning	Secondary meaning(s)
<i>aḍi/aṛi</i>	‘who?’	‘which?’ (Donegan & Stampe 2004)
<i>biṛi/biri</i>	‘what?’	‘who?’ (Dasgupta 1978:40; Donegan & Stampe 2004)
<i>maṇḍi/maṇḍi</i>	‘which?’	‘what?, who?’ (Matson 1964:40), ‘whom?’ (Dasgupta 1978:85; Donegan & Stampe 2004)
<i>nay</i>	‘what?’, only in <i>nay ḍaḍ/ nay ḍuan/ nay ḍaʔn</i> ‘what for?’ (Donegan & Stampe 2004); compare Gorum <i>nay</i> ‘what?’. More commonly ‘what for?, why?’ is given in the sources as <i>bi((r)i)-te</i> , where <i>te</i> is a suffixed or encliticized case marker, depending on one’s analysis.	

well as the marker of a (in)direct animate object (for the meanings of *-a* and *-te* see Dasgupta 1978:59-60).⁵³

Juang (Austro-Asiatic, South Munda; India)

- (77) a. *āi-a kɔte maṇḍi-a?*
 1SG-GEN cloth IPW-LOC
 ‘Where is my cloth?’ (Dasgupta 1978:74)
- b. *aṇ-a paṇṭon maṇḍi te? ere tebulu te*
 1SG-GEN pen IPW LOC DEM/3SG.NON<HUM> table LOC
 ‘Where is my pen? It’s on the table’ (Matson 1964:55)

In this pronominal use, *maṇḍi* appears to be used as a non-selective interrogative pronominal ‘what?’. The question remains, however, whether it can be used as ‘what?’ outside of the construction with a locative marker.

Furthermore, the interrogative *maṇḍi* also looks like a good candidate for qualifying as a ‘who?, what?’ interrogative (besides its selective use), because some of the sources also gloss *maṇḍi* as ‘who?’ and ‘whom?’ (76). Although I found no examples of this use in the sources consulted, there is some indirect evidence supporting these glosses.

First, the link between the selective interrogative ‘which (one)?’ and the human interrogative ‘who?’ is not uncommon cross-linguistically (cf. Section I.2.3). Second, according to Masica (1991:255), in modern Indo-Aryan languages, which are dominant in the region at issue, “the *kaun*, *kon*

⁵³ Dasgupta (1978) describes this and other markers postposed to nominals as “suffixes”. At the same time, in Kharia, which is rather closely related to Juang, these markers are clearly clitics (John Peterson, p.c.).

‘animate/personal’ forms [of the interrogative pronominals, i.e. ‘who?’] generally also can mean ‘which’ when used adjectivally (sometimes with an extension, as in H[indi] *kaun s̄ā*),⁵⁴ while “some languages have *kon* forms (meaning ‘who, which’)⁵⁵ in addition to [the interrogatives meaning only ‘who?’] (e.g., N[epali] *kun* [next to *ko* ‘who?’]).”⁵⁶

Let us now consider the Juang interrogative *biṛi*, which also has both ‘who?’ and ‘what?’ among its glosses in some sources. Usually it is glossed ‘what?’ (78-79), but Dasgupta (1978) and Donegan & Stampe (2004) also provide the gloss ‘who?’.

Juang (Dasgupta 1978:74)

(78) *ape-a iti-a biṛi asi-ke?*
2PL-GEN hand-LOC IPW be-PRS
‘What is in your hand?’

(79) *nii biṛi ni-jime?*
1PL IPW 1PL-eat.FUT
‘What shall we eat?’

Although again no examples of this interrogative being used as ‘who?’ have been found in the sources consulted, there is strong indirect evidence supporting the existence of the gloss ‘who?’.

First, Dasgupta (1978) repeatedly gives the gloss ‘who?’ and even cites a case marked form *biṛi-te* ‘whom? (SG or PL)’ (for *-te* see above). Second, another South Munda language, Kharia, has an apparently cognate form *ber/berhar/behār* meaning ‘who?’ (Donegan & Stampe 2004). Grierson (1906:IV:194) gives the interrogative ‘who?’ in Kharia basically as *ber* (80a), next to a few optional forms of it inflected for person-number, such as *ber-(j)ār* 1DU, *ber-hār* 2DU (80b),

⁵⁴ The *s̄ā* in *kaun s̄ā* ‘which particular one? (out of a number)’ is a “suffix” meaning ‘very, intensely’ (cf. Platts 1884).

⁵⁵ In some languages, this interrogative can be used only attributively, as in the case of Kharia Thar (an Indo-Aryan language, not Munda) *kon/kṛṇ* ‘which [N]?’ (Dasgupta 1978:122).

⁵⁶ For instance, Assamese, as described by Babakaev (1978), has *ki* ‘what?’, *kih* ‘which [N]?, what [N]?’ and *kon* ‘who?, which one? (person)’ but also a selective ‘which [N]? (person or thing)’, as in *kon mānuh* ‘which man? (out of a number)’ and *kon kitāp* ‘which book? (out of a number)’ (1978:75-76). It seems that *kaun/kon* interrogative can be used as ‘which one?’ for things as well, only in the languages where it takes on a postposed emphatic marker in its attributive use (like in Hindi *kaun s̄ā*). The *kaun/kon* interrogative forms in modern Indo-Aryan languages go back to an earlier a combination of *kaḥ*, NOM.M.SG form of the general interrogative pronominal root ‘who?, what?, which (one)?’, and *punar* ‘again, also, moreover, but, however’ (Turner 1966:127, 469).

ber-kiār 3DU.⁵⁷ According to John Peterson (p.c.), in his Kharia data the forms *ber*, *berhar*, and *behar* are used interchangeably without any difference in meaning. Still, Grierson's forms look plausible because the dual in Kharia is also used as honorific (John Peterson, p.c.). That is, we may hypothesize that in modern Kharia the originally dual form *berhar* has become lexicalized as a regular interrogative 'who?'.⁵⁸

Kharia (Austro-Asiatic, South Munda; India)

(80) a. *ber-a koṇḍu?*

IPW-GEN boy

'Whose boy [comes behind you]?' (Grierson 1906:IV:274)

b. *amār ber-hār heke-bār?*

2DU/2HON IPW-2DU/2HON be-2DU/2HON

'Who are you two?' (Grierson 1906:IV:194), or probably 'Who are you (honorific, and irrespective of number)?'

That *biṛi* means 'what?' and apparently marginally 'who?' in Juang, but its cognate in Kharia means 'who?' can be explained rather easily if we assume that this interrogative used to be a selective interrogative pronominal 'which one?' indifferent to the person/thing distinction to 'who?' and 'what?' but was later extended to non-selective contexts. Interestingly, the distribution of the human and non-human meanings of the cognates of *biṛi* in South Munda languages, 'who?' in Kharia but mostly 'what?' in Juang, happens to correlate nicely with the distribution of the meanings of the *kaun/kon* interrogatives within Indo-Aryan, which have been mentioned above. Recall that in modern Indo-Aryan languages there is a strong formal similarity, often even identity between the interrogatives meaning 'who?', 'which [N]? (person or thing)' and 'which one? (person or thing)', the basic form for such interrogatives being *kaun/kon* (or the like). At the same time, in Oriya, an Indo-Aryan language spoken in the region encircling the Juang territory, the interrogative *kəṇḍə*, cognate to *kaun/kon*, means 'what?' and not 'who?' (cf. Masica 1991:255).

Having examined the patterns of use of the Juang interrogatives *biṛi* and *mandi*, let us now consider their possible etymologies. To begin with, note that it seems possible to divide the three Juang interrogatives *biṛi*, *mandi* and *aḍi/ari* in two parts, with the final element *-ḍi/-ṛi* shared by all of them (cf. Donegan & Stampe 2004). The element *-ḍi/-ṛi* is likely to have functioned before as an (attributive) interrogative.⁵⁸ Alternatively, it might have been a kind of focus

⁵⁷ According to John Peterson's data (p.c.) the dual/honorific markers on predicates and personal pronouns in Kharia are =*naṇ* 1DU/HON.INCL, =*jar* 1DU/HON.EXCL, =*bar* 2DU/HON and =*kiyar* 3DU/HON.

⁵⁸ Consider, for instance, a Proto Mon-Khmer *[?]*ci?* 'relative/interrogative pronoun' (Shorto

marker, like Gorum *-di*, or a demonstrative, or still something else. All in all, its exact original function is not so important here.

The initial *a-* in *adi/ari* is probably a deictic or pronominal root (cf. Donegan & Stampe 2004), while *bi-* and *man-* in *biṛi* and *mandi* must have been the nominals meaning ‘place’ or a kind of locative prepositions,⁵⁹ i.e. the interrogatives were originally locative in meaning and were constructed as ‘what place?’ or ‘at/in (the place of) what?’. Note in this respect that this way of encoding the meaning ‘where?’ seems to be particularly common in Mon-Khmer languages (both synchronically and etymologically). For the root *man-* compare Juang *man-(e)* ‘place’ (Donegan & Stampe 2004), the Sora locative case marker *-ə-man* (Ramamurti 1931:20), Wa *maŋ* ‘side, place (combining form used in locative expressions)’ (Austro-Asiatic, Northern Mon-Khmer, Palaungic; Myanmar, China, Thailand; <http://wadict.soas.ac.uk>),⁶⁰ the Temiar preposition *ma* ‘to (direction)’ (Austro-Asiatic, Mon-Khmer, Aslian; Malaysia; Benjamin 1976b:163). For the root *bi-* compare Gorum *ba(ʔ(a))* ‘place’ or ‘at/in (locative), by (instrumental)’ (Donegan & Stampe 2004), Kharia *boʔ* ‘place’ (John Peterson, p.c.), Gutob *bō* ‘where?’ (Grierson 1906:IV:236), Bugan *pəʔ⁵⁵* ‘side’, as in *pəʔ⁵⁵-ɕi⁵⁵* ‘left side’ or *pəʔ⁵⁵-ɕa⁵⁵* ‘right side’ (Austro-Asiatic, Mon-Khmer; China; Li 1996:144), the Semai preposition *bə* in *bə lɔ:ʔ* ‘where?’ (Austro-Asiatic, Mon-Khmer, Aslian; Malaysia; Dentan 2003:77), the Chut “dative prefix” *pa-* ‘to, for, with (somebody)’ (Austro-Asiatic, Mon-Khmer, Viet-Muong; Vietnam; Solntseva & Solntsev 2001:80-81), Khmer *pɿ:* ‘from, about, of’ (Austro-Asiatic, Eastern Mon-Khmer; Cambodia; Gorgoniyev 1966:107) and Chrau *vi* ‘place’ or ‘where?’ (Austro-Asiatic, Eastern Mon-Khmer, South Bahnaric; Vietnam; Thomas 1971:198). If the original vowel of *bi-* was not *i*, than it is likely to have been *a* with a later assimilation *a > e > i* to the vowel *i* of *-di/-ṛi*.

Finally, it can be noted that the development from ‘where?’ to ‘which one? (person or thing)’ is not uncommon in the languages of the world. A further development from a selective interrogative pronominal ‘which one?’ indifferent to the human/thing distinction to ‘who?’ and ‘what?’ is also more easy to

2006, #46; see also footnote 37 in Section III.4.1.3) or various *ci* (or the like) ‘what?’ interrogatives in North Munda languages (Donegan & Stampe 2004). However, the North Munda forms may also be loans from Indo-Aryan, where the interrogative ‘what?’ is often *ki* or the like.

⁵⁹ Although unlike Mon-Khmer languages, Munda languages are generally head-final, there are some elements of head initial syntax as well. For instance, in Gorum the oblique case is marked on pronominals by a preposed marker *e(n)=*, as in *e=moɣ* ‘whom? (OBJ), to whom?’ (Donegan & Stampe 2004).

⁶⁰ For instance, *maŋ liuŋ* ‘above’, *maŋ siuŋ* ‘below, underneath’, *maŋ tiuŋ* ‘over there’ (<http://wadict.soas.ac.uk>).

conceive than a direct development from ‘who?’ to ‘what?’ or vice versa.

4.2 Austronesian languages

4.2.1 Introduction

Austronesian languages are spoken in a vast region covering Taiwan, Philippines, continental South-East Asia, Indonesia, Oceania and Madagascar. Traditionally, the major divide within the Austronesian family is made between the Austronesian languages spoken on Taiwan (commonly believed to be the Proto Austronesian homeland), the so-called Formosan languages, and the rest, the Malayo-Polynesian languages, sometimes also referred to as Extra-Formosan. The latter languages are often conveniently subdivided into Western Malayo-Polynesian and Central-Eastern Malayo-Polynesian. The distribution of the Austronesian languages and their major subdivisions as presented above are illustrated on Map 7.

Map 7. Distribution of Austronesian languages⁶¹



The Austronesian languages of Taiwan, although usually summarily referred to as Formosan, do not seem to form a unity. The existing classifications suggest from two to nine independent subgroups, each of the same time-depth or older than Malayo-Polynesian.⁶²

⁶¹ Austronesian languages: major divisions of Austronesian languages. [Map/Still], *Encyclopædia Britannica Online*: <http://www.search.eb.com/eb/art-2108>, retrieved 14.03.2007

⁶² See, for instance, Li (2006) for a brief overview of various classifications of the Austronesian languages of Taiwan.

Languages with interrogative pronominals used indiscriminately as ‘who?’ or ‘what?’ appear to exist both among Formosan and Malayo-Polynesian languages, including Western and Central-Eastern subdivisions of the latter. I have checked some 450 Austronesian languages.⁶³ Thus, such languages have been found among the Northern Luzon languages (Section III.4.2.2), the Lampungic languages (Section III.4.2.3), the Choiseul languages (Section III.4.2.4), the South Huon Gulf languages (Section III.4.2.5), and the Tsouic languages (Section III.4.2.6). Tsouic is a Formosan family, while the rest are lower-level subdivisions of Malayo-Polynesian.

Numerically, the highest concentration of languages with ‘who?, what?’ interrogatives is found among the Northern Luzon languages, while other groups at issue appear to have just a few such languages, one or two on average. Northern Luzon languages also differ from the other groups in the ways their ‘who?, what?’ interrogatives have come to existence. Thus, in the other groups the ‘who?, what?’ interrogatives mostly appear to go back to an interrogative construction modelled on the structure [‘what?’ + a demonstrative or personal pronominal] and literally meaning ‘what one?’, i.e. ‘which one? (person or thing)’, with a later extension to non-selective contexts both as ‘who?’ and ‘what?’ and in some case with an accompanying loss of the original interrogative element. In the Northern Luzon languages, the ‘who?, what?’ interrogatives (i) represent reflexes of the word meaning ‘name’ through a conventionalization of questions like *Its name?*, *Your name?*, etc., (ii) are due to the peculiarities of the semantics of the noun phrase marker **si* and its reflexes, which become fossilized on these interrogatives. In the Northern Luzon language Isneg and the Choiseul language Sissinga, yet another scenario may be possible. Thus, the ‘who?, what?’ interrogative may also result from a semantic shift of ‘where?’ to ‘which one? (person or thing)’ and a subsequent extension of the latter to non-selective contexts as ‘who?’ or ‘what?’.

Note that two of the three (or four) ways of development of the ‘who?, what?’ interrogatives mentioned above imply a phrasal origin of these interrogatives. It is in all probability no coincidence either that for instance in Blust *et al.* (2006) no monomorphemic interrogative pronominal is reconstructed for Proto Austronesian. Instead, Blust *et al.* (2006) suggest the reconstruction of **n-anu* ‘what?’ and **si-ima* ‘who?’. As will be discussed in Section III.4.2.6.1.1,

⁶³ This number includes approximately 430 Austronesian idioms from (Blust *et al.* 2006, accessed on 20.11.2006) and some 20 additional languages. I have counted the languages in Blust *et al.* (2006) in the following way: (i) the reconstructed forms have not been counted; (ii) a given language has been counted only if there are data both on ‘who?’ and ‘what?’, (iii) I have counted multiple entries for the same language only once (i.e., if for instance, there are two ‘what?’ entries and three ‘who?’ entries for a given language, this still counts as one language).

the latter reconstruction includes the so-called “personal” noun phrase markers, **si* and/or **(?)i*,⁶⁴ and the interrogative root proper here is in all probability **ma* ‘what?’. At the same time, the reconstructed form **n-anu* strongly resembles another interrogative reconstructed in Blust *et al.* (2006) for Proto Austronesian, **i-nu* ‘where?’, where the initial *i* is in all probability a (locative) preposition, which is not uncommon in, for instance, the Formosan Austronesian languages. The initial *na-* or *a-* of **n-anu* may also be cognate to similar demonstratives and/or noun phrase markers. The root **nu* itself may have been polysemous between ‘where?’ and ‘which (one)?’, as is also often the case in many modern Austronesian languages (cf. Section III.4.2.2.2). Alternatively, as suggested by Reid (1981), Proto Austronesian **nu* should be reconstructed as ‘thing’.⁶⁵ The forms like Dempwolff’s (1938) **anu* and Dahl’s (1976) **an₁u* ‘somebody, something’, and apparently also Blust *et al.*’s (2006) **n-anu* ‘what?’ and **i-nu* ‘where?’, are then based on a combination of **nu* ‘thing’ with some preposed noun phrase marker. This appears to imply that **nu* developed its interrogative pronominal function only at some post Proto Austronesian stage, although apparently a rather early one, while **ma* ‘what?’ remains the only monomorphemic interrogative pronominal that should be reconstructed for Proto Austronesian. By way of conclusion, it is most probable that besides the two monomorphemic interrogative pronominal roots, Proto Austronesian and its early daughters also resorted to constructions based on them to express various interrogative pronominal meanings in a more explicit or emphasized way.⁶⁶

⁶⁴ Note that of the two markers, **si* and **(?)i*, the second one is probably the oldest (cf. Reid 1979:11-16; see also Section III.4.2.2.2).

⁶⁵ It is interesting to mention that actually, Reid (1981) proposes the Proto Austronesian reconstruction **nu* ‘thing’ to account for the widespread Austronesian genitive marker **nu* and only uses the interrogatives (and indefinites) based on **nu* as a piece of evidence for the etymology of the genitive.

⁶⁶ Given that since Schmidt’s (1906) paper there have been attempts to link Austronesian to Mon-Khmer (and Austro-Asiatic in general) in a greater Austric macro-family (see, for instance, Sagart *et al.* 2005 for recent papers and discussion), it may be interesting to mention that both in the forms of its interrogative pronominals and in its strategies for creating new constructional interrogative pronominals, Proto Austronesian appears to be remarkably similar to Proto Mon-Khmer (i.e., non-Munda) Austro-Asiatic. Thus, as far as the forms are concerned, compare Proto Mon-Khmer **m[o]ʔ*/**m[o]h* ‘what?’ (Shorto 2006, #136) vs. Proto Austronesian **ma* ‘what?’, as well as Proto Mon-Khmer **nɔʔ*/**nɔh* ‘this’ (Shorto 2006, #92), which in several daughter branches resulted in **nɔʔ* ‘which?, what?’ (Shorto 2006, #92a), vs. Proto Austronesian **nu* (presumably) ‘thing’, which also resulted in interrogatives in many daughter languages. For more details and discussion of the Austro-Asiatic data see Section III.4.1.1, footnote 3, and subsections of Section III.4.1.

In what follows, I will further elaborate on the Austronesian languages with ‘who?, what?’ interrogatives. I will first discuss the Malayo-Polynesian Austronesian languages in the following order: Northern Luzon (Section III.4.2.2), Lampungic (Section III.4.2.3), Choiseul (Section III.4.2.4), and South Huon Gulf (Section III.4.2.5). The section will be concluded by examining the Tsouic languages, a Formosan Austronesian group (Section III.4.2.6).

4.2.2 Northern Luzon languages⁶⁷

Northern Luzon languages, earlier known as Cordilleran, are spoken in the northern parts of Luzon Island and belong to the Northern Philippine subdivision of the Malayo-Polynesian branch of the Austronesian family. According to the Ethnologue, the Northern Luzon group consists of four branches Figure 6.

Figure 6. Northern Luzon languages according to the Ethnologue

- Arta
- Ilocano
- South-Central Cordilleran
- Northern Cordilleran

The first two branches contain only one language each, whereas the last two have some thirty and twenty languages respectively. Except Ilocano, most Northern Luzon languages are numerically rather small. Yet, sometimes they provide examples of an astonishing degree of small-scale variation, particularly in their systems of noun phrase markers and interrogative pronominals.

⁶⁷ I am grateful to Lawrence Reid for helpful comments and interesting discussion, as well as for sharing with me his Bontok data. I would also like to thank him for the kind permission to cite the unfinished (2006) version of (Reid, forthcoming). The (2006) version does not seem to differ considerably from the final version as far as the pieces of data and analysis cited here are concerned. It should be mentioned, however, that I took the liberty to slightly modify the presentation of the data cited from or via this and other of Reid’s papers. The two major points where the presentation here differs from that of Reid are as follows. First, I use the terms ergative/genitive and absolutive, where Reid would use (only) genitive and nominative respectively. Although Reid considers the Philippine languages discussed to be “ergative”, he does not use the term absolutive because in his view the term nominative “captures more linguistic generalities”, while he prefers the term genitive to ergative because “the forms which mark the actor in a transitive clause are identical to those which mark nominal possessors within a noun phrase” (2006, forthcoming). The second point concerns my use of the zero mark \emptyset , mostly for the absolutive function (Reid’s nominative). Reid himself (p.c.) disapproves the use of a zero mark because “it implies a zero morpheme, that [he] do[es]n’t believe in”.

Languages with interrogative pronominals used indiscriminately as ‘who?’ or ‘what?’ are found primarily in the South-Central Cordilleran branch (Section III.4.2.2.1). Furthermore, there is at least one such language in the Northern Cordilleran branch (Section III.4.2.2.2). As far as the origins of the lack of differentiation between ‘who?’ and ‘what?’ are concerned, the Northern Luzon languages with ‘who?, what?’ interrogatives can be divided in two to three groups. The first two groups comprise several South-Central Cordilleran languages. The ‘who?, what?’ interrogatives in these languages either represent reflexes of the word meaning ‘name’ or are due to the peculiarities of their systems of noun phrase markers and particularly to the semantics of the noun phrase marker **si* and its reflexes. The third group may comprise at least one Northern Cordilleran language. The ‘who?, what?’ interrogative in this language may result from a semantic shift of ‘where?’ to ‘which one? (person or thing)’ and the subsequent extension of the latter to non-selective contexts as ‘who?’ or ‘what?’. Alternatively, the latter Northern Cordilleran language may be similar to the South-Central Cordilleran languages of the second group, with only difference that the noun phrase marker involved is not **si* but its predecessor **ʔi*.

4.2.2.1 South-Central Cordilleran languages⁶⁸

The South-Central Cordilleran languages of Northern Luzon are spoken in a rather compact inland mountainous area in the northern part of Luzon Island. The internal classification of South-Central Cordilleran languages as found in the Ethnologue is provided in Figure 7. Languages with interrogative pronominals used indiscriminately as ‘who?’ or ‘what?’ are found in Central (Section III.4.2.2.1.3) and Southern Cordilleran subgroups (Section III.4.2.2.1.4). I do not have information on the Alta languages. It is difficult to say with certainty how many South-Central Cordilleran idioms with ‘who?, what?’ interrogatives there are exactly because there appears to be a good deal of variation in the inventories of interrogative pronominals, sometimes even within what is usually considered to be a dialect of a language. The languages I found to have ‘who?, what?’ interrogatives include several varieties of Kalinga (Section III.4.2.2.1.3.1), such as Upper and Lower Tanudan Kalinga (Sections III.4.2.2.1.3.1.1-III.4.2.2.1.3.1.2), a variety of Limos Kalinga (Section III.4.2.2.1.3.1.3), and

⁶⁸ I would like to thank Sherri Brainard for sharing with me her data on Upper Tanudan Kalinga, Limos Kalinga and Karao, helpful comments and interesting discussion. I am also grateful to her for bringing me in touch with her colleagues working on Lower Tanudan Kalinga and Lagawe Ifugao.

*Figure 7. South-Central Cordilleran languages according to the Ethnologue***Alta**

- Southern Alta
- Northern Alta

Central Cordilleran

- Isinai
- Kalinga-Itneg languages
- Nuclear Cordilleran
 - Balangao
 - Bontok-Kankanay languages
 - Ifugao languages

Southern Cordilleran

- Ilongot
- Pangasinic
 - Pangasinan
 - Benguet: Kallahan languages, Ibaloi-Karao, I-Wak

Guinaang Kalinga (Section III.4.2.2.1.3.1.4).⁶⁹ Furthermore, several Nuclear Cordilleran languages of this kind have been identified (Section III.4.2.2.1.3.2), such as Lagawe and Bayninan Ifugao (Section III.4.2.2.1.3.2.1), Northern Kankanay (Section III.4.2.2.1.3.2.2), and Natonin Balangao (Section III.4.2.2.1.3.2.3). Two Southern Cordilleran languages with a ‘who?, what?’ interrogative are Kayapa Kallahan (proper) and Keley-i Kallahan (Section III.4.2.2.1.4).

It is possible to divide the South-Central Cordilleran languages with ‘who?, what?’ interrogatives in two groups according to the origins of these interrogatives. The first group comprises Upper and Lower Tanudan Kalinga, Lagawe and Bayninan Ifugao and a variety of Limos Kalinga. The second group contains Guinaang Kalinga, Northern Kankanay, Natonin Balangao, and Kayapa and Keley-i Kallahan. Before presenting the data of particular Central (Section III.4.2.2.1.3) and Southern (Section III.4.2.2.1.4) Cordilleran languages, I will further elucidate on the general mechanisms that I believe to account for the existence of the ‘who?, what?’ interrogatives in the first group of languages (Section III.4.2.2.1.1) and in the second group of languages (Section III.4.2.2.1.2).

⁶⁹ According to Sherri Brainard (p.c.), in all probability most varieties of Kalinga do not differentiate between ‘who?’ and ‘what?’.

4.2.2.1.1 The ‘name’-based interrogatives

The ‘who?, what?’ interrogatives in Upper and Lower Tanudan Kalinga, Lagawe and Bayninan Ifugao and a variety of Limos Kalinga represent reflexes of the word meaning ‘name’.⁷⁰ For Proto South-Central Cordilleran, the root ‘name’ is reconstructed as **ɲadan/*ɲadən* (Reid 2006, with a reference to a personal communication by Ronald S. Himes for the reconstruction of the doublet form) and for Proto Austronesian as **ɲajan* (Blust *et al.* 2006).

The development of a nominal meaning ‘name’ to an interrogative pronominal must have proceeded through a conventionalization of questions like *Its name?*, *Your name?*, *The name of the one that fell?* (person or thing), etc. Such questions are possible even in English, although they are pragmatically rather marked. However, there are quite a few languages in the world which do not normally use a question word when inquiring about names. For, instance this appears to be typical for West Africa, which may be illustrated with (81) from Bamana. A longer form of (81) using an interrogative *dì* ‘how?’ is also possible, but by no means necessary.

Bamana (Niger-Congo, Western Mande; Mali)

- (81) *í tɔ́gɔ́ (bé dì)?*
 2SG name-ART be how
 ‘What’s your name?’ (lit.: ‘Your name (is how)?’)

It is worth mentioning here that for a ‘what?’ or ‘who?, what?’ interrogative to develop from a word meaning ‘name’, another necessary requirement needs to be met. The word meaning ‘name’ must have a broad intension, in the sense that like in English it must be possible to speak about *the name of a person*, *the name of a mountain*, *the name of a tree*, *the name of this thing*, etc.⁷¹

Examples of a ‘who?, what?’ based on the word meaning ‘name’ will be discussed in more detail in the sections dedicated to the particular languages, cf. Sections III.4.2.2.1.3.1.1-III.4.2.2.1.3.1.3 and III.4.2.2.1.3.2.1-III.4.2.2.1.3.2.2.

⁷⁰ I owe this hypothesis to Sherri Brainard (p.c.). See also Reid (1979:9-10), who seems to have been the first to suggest in print the development from ‘name’ to an interrogative pronominal for the Philippine languages. However, since he did not consider languages such as Limos Kalinga, he has suggested the development from ‘name’ only for the non-human interrogative ‘what?’.

⁷¹ Remarkably, even closely related languages may differ considerably in this respect. Thus, although many South-Central Cordilleran languages appear to use (or to have used) their words for ‘name’ in such a “loose” way, as can be deduced from the development of interrogative pronominals from the words meaning ‘name’ in these languages, there are also many that do not allow this. For instance, according to Lawrence Reid (p.c.), in Bontok the word *ngachan* ‘name’ can be used to speak about names of persons or places but not about names of trees.

Remarkably, both among South-Central Cordilleran languages and outside of this group, there are also numerous examples of the word ‘name’ being involved in the formation of dedicated ‘who?’ and ‘what?’ interrogatives, which provide additional support to the ‘name’ hypothesis.

For instance, the Paril dialect of Kalamansig Cotobato Manobo, which is a Southern Philippine language spoken in the south of Mindanao Island far away from Northern Luzon, has *ngadan* meaning both ‘name’ and ‘what?’. The meaning ‘who?’ is expressed with a phrase *ngadan di ?itaw*, where *?itaw* is ‘person’ and *di* is a noun phrase marker (Reid 1971:107, 116, 161, 163). That is, ‘who?’ must literally mean something like ‘name/what of person?’, i.e. ‘what person?’. In Ibaloi, a Southern Cordilleran language, “the common word for ‘what?’ is *nġaram-to* ‘its name?’”,⁷² with a contracted form *ngan-to* (Blake 1906:362-363). Karao, another Southern Cordilleran language, has three special interrogative expressions *nganiyay* ‘what is this (near me)?’, *nganithan* ‘what is this (near you)?’ and *nganiwan* ‘what is that? (at a distance both from me and you)?’ (Sherri Brainard, p.c.).⁷³ The three interrogatives clearly result from the fusion of *ngaran* ‘name’ followed by a genitive marker *ni* and a demonstrative root *yay* ‘this (near me)’, *than/tan* ‘this (near you)’ and *man/wan* ‘that (at a distance both from me and you)’. That is, literally the whole construction used to mean ‘the name of this/that?...’. Similarly, another Southern Cordilleran language Kakidugeen Ilongot has *ngadi:n ni-ma* ‘what?’ (Reid 1971:161), which appears to have been constructed following the same principle as Karao *nganiwan*, i.e. [‘name’ + ‘of’ GEN + ‘that’]. Interestingly, the interrogative ‘who?’ *ka-tu?u* in Kakidugeen Ilongot includes the word *tu?u* ‘person’ (Reid 1971:116, 163), while the initial *ka* appears to be a reflex of a (human) “dative” marker **ka* (cf. Reid 2006).⁷⁴ That is, literally *ka-tu?u* used to mean something like ‘of person?’. In all probability, this interrogative represents a reduction of an original construction **ngadi:n ka-tu?u* ‘name of person?’, similar to the Paril Kalamansig Cotobato Manobo expression for ‘who?’ *ngadan di ?itaw* mentioned above.

See also Reid (1979:9-10) for some further examples of the development of the word for ‘name’ to an interrogative pronominal in Philippine languages. Interestingly, Reid (1979:9) points out that “the use of the word for ‘name’ as an interrogative may be a development that predates Proto-Philippines”, given that “Dempwolff (1938) cites Ngadju-Dayak [(Malayo-Polynesian, West Barito;

⁷² According to Lawrence Reid (p.c.), in this form the letter *m* should be *n*.

⁷³ The regular interrogative pronominals of Karao are *nengo* ‘what?’ and *siya* ‘who?’ (Sherri Brainard, p.c.).

⁷⁴ Reid (2006) uses the term “dative” because of “its general use for phrases expressing [human] recipients”.

Indonesia, Kalimantan)] *'ara* ‘name’, as well as *n-ara-i* ‘what’, as evidence for his reconstruction of **[ʔag'an* ‘name’”.

4.2.2.1.2 The **si*-initial interrogatives

The ‘who?, what?’ interrogatives of Guinaang Kalinga, Northern Kankanay, Natonin Balangao, and Kayapa and Keley-i Kallahan are all based on the same model: a frozen reflex of a noun phrase marker **si*, which is typically *si* or *hi*, followed by a reflex of one of the two Proto Malayo-Polynesian ‘what?’ roots, **-anu* or **apa* (cf. Blust *et al.* 2006), e.g. Guinaang Kalinga *sinu* or Kayapa Kallahan and Keley-i Kallahan *hipa(?)*.

Noun phrase markers are typically monosyllabic forms introducing a noun phrase in many Austronesian, and particularly Philippine languages. In Northern Luzon languages, they also often become encliticized to the preceding element, although functionally they relate to the element that follows (Reid 2002, 2006). These markers are used to encode a wide variety of semantic and functional features, most importantly related to case and various kinds of deixis and reference disambiguation. There is a vast literature on the issue and many different terms have been proposed to describe these elements. An overview can be found, for instance, in Reid (2002). According to the analysis in Reid (2000, 2002, 2006), the noun phrase markers can be subdivided into two categories, case prepositions and nominal specifiers.⁷⁵ It is claimed that historically the noun phrase markers originate in demonstratives (they are often still homophonous to them) and that the case markers have developed from demonstratives through a nominal specifier stage. A detailed study of the evolution of the system of noun phrase markers in Northern Luzon languages in general and South-Central Cordilleran languages in particular can be found in Reid (2006).

Given that historically the ‘who?, what?’ interrogatives in the South-Central Cordilleran languages in question are based on the structure [nominal specifier **si* + **anu/*apa* ‘what?’], it is reasonable to assume that it is the nominal specifier **si* that accounts for this lack of differentiation between ‘who?’ and ‘what?’. Therefore, in the rest of the present section I will first examine this nominal specifier in more detail and then advance a hypothesis regarding its earlier functions, which may help us to account for the lack of differentiation between ‘who?’ and ‘what?’ in the interrogatives at issue.

In other Malayo-Polynesian languages the interrogatives of the structure [nominal specifier **si* + **anu/*apa* ‘what?’], normally mean ‘who?’, e.g. Tagalog *sino* ‘who?’ vs. *ano* ‘what?’ (Austronesian, Malayo-Polynesian, Meso

⁷⁵ Nominal specifiers are “the heads of their phrases” and can be largely glossed as ‘the one’ (Reid 2002:295).

Philippine; Blust *et al.* 2006) or Indonesian *si-apa* ‘who?’ vs. *apa* ‘what?’ (Austronesian, Malayo-Polynesian, Malayic; Blust *et al.* 2006). Besides their use in ‘who?’ interrogatives, the reflexes of the noun phrase marker **si* in many Malayo-Polynesian languages are often confined to the marking of personal proper names in certain syntactic contexts. Not surprisingly, the noun phrase marker *si* is usually described as a “personal” marker, which roughly speaking is a short-cut for saying a personal proper name marker. However, in the languages of Northern Luzon the noun phrase marker *si* shows up in a much wider variety of contexts. To get an idea of the range of functions that *si* may have in Northern Luzon languages, consider for instance the system of case markers⁷⁶ and nominal specifiers of Guinaang Kalinga (Table 2 and Table 3).

Table 2. Guinaang Kalinga case markers (based on Reid’s 2006 adaptation of Gieser 1963:50)

TOP/PRED		ABS	GEN/ERG	OBL	LOC
<i>si</i>	after V:	Ø	= <i>n</i>	= <i>s/</i> = <i>ʔ</i>	= <i>d/</i> = <i>ʔ</i>
	after C:		Ø	<i>si</i>	<i>ʔud/ ʔad*</i>

*This form occurs in “the dialect of Kalinga spoken in Lubuagan”

Table 3. Guinaang Kalinga nominal specifiers (Reid’s 2006 adaptation of Gieser 1963:50)

Indefinite		<i>di/ si</i>
	default	<i>nan/ = n</i>
Definite	-PST	<i>din/ = n</i>
	+PST	<i>dit/ = t</i>

Examples (82-84) illustrate the topic marker, oblique marker and indefinite nominal specifier function of *si* in Guinaang Kalinga. Note that *tagu* ‘person’ in (84) is not a personal proper name, but just a human noun.

⁷⁶ Reid (2006) notes the following about presenting the topic and nominal predicate marker together with the case markers: “Although I do not consider topics and nominal predicates to be ‘case-marked’, the prepositions that introduce them function in a similar way to those that case-mark other NPs, in that they identify the function of the NP, and particularly in the case of topics they serve also to foreground the phrase within a discourse”.

Guinaang Kalinga (Gieser 1963:18, 24, 28, 59, 76-78, 86 cited via Reid 2006)

- (82) *si=n ?anak=a lalaki ?immoy nangayu*
 TOP=NS child=LIG male went collected.wood
 ‘As for the boy, he went to get wood.’
- (83) *?umiwas=?ami si ?iwoy si=t gattok=a saklot=na*
 shave=ABS.1PL OBL rattan OBL=NS permanent=LIG tie=GEN.3SG
 ‘We shave rattan for its permanent ties.’
- (84) *patoyon=da Ø si tagu*
 kill=[ERG]3PL ABS NS person
 ‘They kill a person.’

Unfortunately, I do not have Guinaang Kalinga examples of the nominal specifier *si* being used with non-human nominals. Instead, we may consider some examples from Guinaang Bontok, another Central Cordilleran language but of the Nuclear Cordilleran subgroup. According to Reid (2006), Guinaang Bontok *si* can be both a personal nominal specifier (85) and an indefinite (or non-specific) nominal specifier with common nouns (86-87).

Guinaang Bontok (Reid 2006)

- (85) *kinmaan Ø si Takcheg*
 departed ABS NS_{+PERSONAL} PROP
 ‘Takcheg departed.’
- (86) *chaan omeý Ø si esa=y minotos...*
 NEG.PST go ABS NS_{-PERSONAL} one=LIG minute
 ‘A minute had not yet gone by...’
- (87) *inára=n si ótot Ø na=ófi*
 got=ERG NS_{-PERSONAL} rat ABS NS_{-PERSONAL}=sweet.potato
 ‘Rats/A rat/Some rats ate the sweet potato.’

However, as a personal nominal specifier Guinaang Bontok *si* occurs only when the noun functions as a nominal predicate, topic or absolutive and only when this noun is a singular human proper name or kinship term of address. In this respect, compare also (88-89). Furthermore, Guinaang Bontok *si* is an oblique case marker and a “conjunction conjoining comparable groups of non-specific human nouns” (Reid 2006).

Guinaang Bontok (Reid 2006)

- (88) *kinmaan Ø nan laráki*
 departed ABS NS_{-PERSONAL} man
 ‘The man departed.’

- (89) a. *sined* Ø *nan* *laráki* Ø *saken*
 waited ERG NS-PERSONAL man ABS 1SG
 ‘The man waited for me.’⁷⁷
- b. *sined* Ø *Takcheg* Ø *saken*
 waited ERG PROP ABS 1SG
 ‘Takcheg waited for me.’

According to Reid (2006), although in most present-day Northern Luzon languages the category “personal nouns” includes only “the names of people and names given to pet animals and animal participants in stories”, in all probability it also used to include “kinship terms of address, as well as titles” (Reid 2006), as it is still the case in Guinaang Bontok. Furthermore, he suggests that in Proto North Central Cordilleran⁷⁸ “the basic semantic distinctions marked by NSs [i.e., nominal specifiers] were personal vs. non-personal”.⁷⁹

However, even if this was indeed the case, it seems that the distinction was rather blurred in the sense that the noun phrase marker *si* was not only a singular personal nominal specifier but also a non-personal nominal specifier of some kind (as well as perhaps an oblique case marker).⁸⁰ It must have been later that the nominal specifier *si* has become restricted in most languages to use with personal nouns only. Presumably, as far as the personal vs. non-personal distinction is concerned, the system of noun phrase markers was rather like the one of Guinaang Bontok, mentioned above. The most important piece of evidence for this comes from South-Central Cordilleran languages such as Keley-i Kallahan (Section III.4.2.2.1.4).

Keley-i Kallahan has a ‘who?, what?’ interrogative *hipa(?)*, showing the original structure [nominal specifier **si* + **apa* ‘what?’] and at the same time a (singular) personal nominal specifier *hi*, which is a regular reflex of **si*. The system of Keley-i Kallahan case markers and nominal specifiers is summarized in Table 4 and Table 5 respectively.

⁷⁷ Note that the absolutive is always unmarked, whereas the genitive/ergative becomes unmarked only when the preceding word ends in a consonant. Otherwise, it is marked by an enclitic =*n* on the preceding word, as in (87).

⁷⁸ That is, the common ancestor of the Nuclear Cordilleran and Kalinga-Itneg languages.

⁷⁹ Consider also similar statements in Reid & Liao (2004:469) for a wider range of Austronesian languages.

⁸⁰ Remarkably, **ʔi*, the predecessor of **si* in the parent of the Philippine languages, is also reconstructible as both a common and a personal nominal specifier and at the same time as a locative preposition (locative case marker) (cf. Section III.4.2.2.2).

Table 4. *Keley-i Kallahan case markers*
(based on Reid's 2006 adaptation of Hohulin & Hale 1977)

TOP		ABS	GEN/ERG/OBL	LOC
<i>(ya)</i>	after V:	\emptyset	$=n$	$=d$
	after C:		<i>ni/</i> \emptyset	<i>di</i>

Table 5. *Keley-i Kallahan nominal specifiers*
(Reid's 2006 adaptation of Hohulin & Hale 1977)

-PERSONAL	default	<i>hu</i>
	proximate	<i>eya</i>
	medial	<i>etan</i>
	distal	<i>eman</i>
+PERSONAL	-PL	after V: $=h$ after C: <i>hi</i> , ERG: <i>nan/</i> \emptyset
	+PL	<i>di</i>

If we assume that **si* was originally a personal nominal specifier, it will be difficult to explain how it could have become frozen on an interrogative meaning not only 'who?' but also 'what?', and at the same time kept on being a personal nominal specifier. Alternatively, if we assume that **si* was originally a nominal specifier that could occur both on common and personal nominals (even though probably more frequently with the latter), it is very easy to account both for the development of the 'who?, what?' interrogatives based on the model [nominal specifier **si* + **anu/*apa* 'what?'] in various South-Central Cordilleran languages *and* for the later evolution of the nominal specifier **si* to a dedicated personal nominal specifier in most Austronesian languages where its reflexes are found to be used as a nominal specifier. In particular, my hypothesis is that earlier the nominal specifier **si* was (optionally?) used to mark (singular) nominals in a topic, predicate or absolutive function⁸¹ that "are not in need of referential disambiguation" (cf. Van de Velde 2006 on this term), either because they are already inherently definite or because they cannot be disambiguated.

Among inherently definite nominals one typically finds proper names, personal pronouns and kinship terms, which provide a strong association with the feature [+human], but also demonstratives and nominals functioning as topics. Remarkably, in many South-Central Cordilleran languages **si* can be found in all

⁸¹ That is, functions which often remain morphologically unmarked in South-Central Cordilleran languages (cf. Reid 2006). In this respect, compare also the Chut prefix *a-* presented below in the present section (Austro-Asiatic, Mon-Khmer, Viet-Muong, Vietnam).

these contexts. Recall, for instance, the use of *si* as a topic or nominal predicate marker and a non-specific nominal specifier in Guinaang Kalinga and its use as a (singular) personal nominal specifier and a non-specific common nominal specifier in Guinaang Bontok. In Lagawe Ifugao (Anne West, p.c.), *hi*=*h*, which is a reflex of **si*, marks nominal predicates, topics or absolutes when these are personal proper names, kinship terms (90) or demonstratives (91), while with other types of nominals the nominal specifiers *di*=*y* or *nan* need to be used.⁸²

Lagawe Ifugao (Anne West, p.c.)

- (90) \emptyset *hi* *aydu*-?
 ABS NS sibling.in.law-GEN.1SG
 '[A: What is Mary's relationship to you? B:] (She is) my sister-in-law.'
- (91) *nganne*=*h* *tuwe*=*n* *kayiw*?
 IPW=[ABS]NS this=LIG tree
 'What is this tree?'

In this respect, compare also the Malayo-Polynesian languages of the Lampungic branch, spoken in the west of Sumatra Island (Indonesia), which show reflexes of "some kind of personal marker" **s(i)*- or **h(i)*- on several personal pronouns and the interrogative 'who?', but also on demonstratives, which are not restricted to human reference (cf. Anderbeck 2006:83).⁸³

As far as the impossibility or unwillingness by the speaker to referentially disambiguate a certain nominal is concerned, some common contexts where this is normally the case may be provided by nominal predicates or non-specific (or indefinite, generic, etc.) nominals. In this respect, recall for instance the nominal predicate marker *si* and the non-specific (or indefinite) nominal specifier *si* of Guinaang Bontok mentioned above.

I owe the aforementioned description "not in need of referential disambiguation" to Van de Velde's (2006) paper on the peculiarities of the

⁸² Note that Anne West herself characterizes all these nominal specifiers as absolute case markers. This is a possible interpretation indeed, but the nominal specifier analysis à la Reid (2006), with unmarked absolutes, topics and nominal predicates, seems to be preferable from a comparative point of view.

⁸³ The reconstruction of the two forms with *i* is mine. Strangely enough, Anderbeck (2006:83) himself reconstructs the first of the two morphemes with a vowel *a*, as **sa*-, and the second as **h*-, despite the existence of such forms as *siapo* 'who?' in the Blambangan Pagar variety of Lampung (Anderbeck 2006:71). I suspect that the form of Anderbeck's reconstruction of the marker in question as **sa*- has been influenced by the existence of a demonstrative *sa* 'this' in some varieties of Lampung (Anderbeck 2006:69). Note in this respect that in the South-Central Cordilleran languages demonstratives of the form *sa* (and the like) and similar nominal specifiers coexist with the nominal specifiers of the form *si* (and the like; cf. Reid 2006).

gender and agreement systems of the Bantu languages of Africa, particularly of the so-called gender G1a and the respective agreement pattern AG1. Remarkably, in some respects the parallel with Bantu may be quite impressive. Thus, although the gender G1 and the respective agreement pattern AG1 are traditionally described as “human” in Bantu studies, in many Bantu languages this “human” agreement has been reported to be used with non-human and even non-animate controllers, such as for instance interjections or infinitives, with generic referents, indefinite referents and for so-called enforced agreement (cf. Section III.2.1.1.2.1). This is reminiscent of the use of the “personal” nominal specifier *sí* with non-human nouns, as well as with topics and nominal predicates, in various South-Central Cordilleran languages. The analysis proposed by Van de Velde (2006) suggests that agreement pattern AG1 in Bantu can mark agreement not only with controllers that belong to gender G1 or have human (or animate) reference, but also those that “are not in need of referential disambiguation”, either because they are already inherently definite or because they cannot be disambiguated.

However, one does not even have to go as far as Africa to find similar nominal markers. Functionally somewhat similar, although admittedly less omnipresent markers appear to exist in various Austro-Asiatic languages (primarily, among what is traditionally referred to as the Mon-Khmer branch of Austro-Asiatic). These markers usually have the form *a-* and the like. Thus, Benjamin (1976b:162) reports that in Temiar (Austro-Asiatic, Mon-Khmer, Central Aslian; Malaysia) “the meaning of most demonstratives may be made more definite by the proclitic /ʔa-/ ‘definer’, which is used with the same function with [...] kin terms and personal names”. According to Kruspe (1999:274-275), in Semelai (Austro-Asiatic, Mon-Khmer, Central Aslian; Malaysia) “the third person free pronouns can be preceded by a prefix ʔa= ‘DET[erminer]’”, the form thus marked “is used for post-verbal arguments encoding absolutive pronominals” and it is described by “speakers [...] as meaning ‘actually’ or ‘in fact him/them’, suggesting its function is to mark the pronominal as referential, specific, etc.”. Wallace (1966:58-60) describes a Katu “prefix” *a-* used to derive demonstrative pronominals from locative adverbs, e.g. *a-đi* ‘this one (here)’ from *đi* ‘here’, and to form focalized (or “emphasized”) forms of personal pronominals (Austro-Asiatic, Eastern Mon-Khmer, East Katuic; Vietnam). In another Katuic language, Ta’oih (Vietnam; Solntseva 1996), the same prefix *a-* appears to occur optionally on personal pronominals, demonstratives, interrogatives, and (indirectly) on kinship terms, but only when these are subjects or objects, that is when they are not headed by a preposition and “not used as an attribute or in a locative function”.⁸⁴ Somewhat similarly in

⁸⁴ See also Reid (1999:13-18) on Ta’oih *a-* and its possible Austronesian parallels.

Chut (Austro-Asiatic, Mon-Khmer, Viet-Muong; Vietnam; Solntseva & Solntsev 2001:72-73, 78), personal pronominals (also with a non-human reference) and the interrogative ‘who?’ may be optionally preceded by a prefix *a-*, but only when they occur in “object referring functions”,⁸⁵ which includes a subject, object and a nominal predicate function and excludes a locative and an attributive function, as well as a dative, marked by a “prefix” *pa-*, and a “possessor” function, often marked by a preposition *kuə*.⁸⁶ Khasi (Austro-Asiatic, Northern Mon-Khmer, Khasian; India) has an “emphatic marker” *ma* which can be used only with personal pronouns and an interrogative pronominal *no*, when these are subjects or nominal predicates (cf. Nagaraja 1985:102; Rabel 1961:68-69; Subbarao & Temsen 2003:201; see also Section III.4.1.2.2.1.1 for examples).

All the aforementioned facts from various Austro-Asiatic languages are reminiscent of the patterns of use of the nominal specifier *si* in South-Central Cordilleran languages. Similarly, *si* normally appears only on nominals with a topic, predicate or absolutive function, which are the functions normally left morphologically unmarked in these languages. Personal names, personal pronominals, kinship terms and demonstratives are also the kinds of nouns that *si* may typically mark.

4.2.2.1.3 Central Cordilleran languages

4.2.2.1.3.1 Kalinga languages

4.2.2.1.3.1.1 Upper Tanudan Kalinga

Upper Tanudan Kalinga is a Kalinga idiom spoken by some 3000 people in the southern end of the Tanudan Valley on Luzon Island. According to Sherri Brainard (p.c.), Upper Tanudan Kalinga uses one interrogative pronominal *ngan* both as ‘who?’ and ‘what?’. She believes that *ngan* ‘who?, what?’ goes back to the word *ngaran* ‘name’, similarly to Limos Kalinga, where according to her data, the same form *ngaran* is used both as ‘name’ and as a ‘who?, what?’ interrogative.

4.2.2.1.3.1.2 Lower Tanudan Kalinga

Lower Tanudan Kalinga is a Kalinga idiom spoken by some 11000 people in a small region on Luzon Island right to the north-east of Upper Tanudan Kalinga. According to Glenn & Jewell Machlan (p.c.), Lower Tanudan Kalinga uses one

⁸⁵ This is a translation of the Russian term *predmetnye funkcii* used by Solntseva & Solntsev (2001:72).

⁸⁶ The preposition *kuə* marks not only possessors, but also some kinds of human oblique arguments.

interrogative pronominal *ngai* both as ‘who?’ and ‘what?’ (92-95).

Lower Tanudan Kalinga (Glenn & Jewell Machlan, p.c.)

- (92) *ngai angtoḃ kan sika?*
 IPW PFV.bite OBL 2SG
 ‘Who/What bit you?’
- (93) *ngai Ø kawadi=n John?*
 IPW ABS work=GEN PROP
 ‘What is John’s work?’
- (94) *ngai Ø ak Juan de la Cruz?*
 IPW ABS NS PROP
 ‘Who is Juan de la Cruz?’
- (95) *ngai=ka?*
 IPW=2SG[ABS]
 ‘Who are you?’

Like Upper Tanudan Kalinga *ngan*, Lower Tanudan Kalinga *ngai* clearly contains a reflex of the word *ngadan* ‘name’. The final *-i* of *ngai* must be a fused reflex of a nominal specifier **di/* =y*, reconstructed by Reid (2006) for Proto North Central Cordilleran⁸⁷ as an “indefinite nominal specifier”. In Northern Luzon languages, noun phrase markers often become encliticized to the preceding element, although functionally they relate to the element that follows, as the genitive marker *=n* in (93) above, which is a clitic form of *ni* as in (96).

Lower Tanudan Kalinga (Glenn & Jewell Machlan, p.c.)

- (96) *ngai Ø ngadan ni anna=e kayu?*
 IPW ABS name GEN DEM=LIG tree
 ‘What is the name of this tree?’

4.2.2.1.3.1.3 Limos Kalinga

Limos Kalinga is a relatively large Kalinga idiom spoken by at least 20000 people in a region right to the north of Lower Tanudan Kalinga. According to Sherri Brainard (p.c.), Limos Kalinga uses the same form *ngaran* both as the word ‘name’ and an interrogative ‘who?, what?’.

The only existing published description of Limos Kalinga grammar, Ferreirinho (1993), describing the variety spoken in Asibanglan, contains somewhat different forms. Unfortunately, Ferreirinho (1993) does not provide any description of the interrogative pronominal system of the language. The few examples with interrogative pronominals that can be found in this source are

⁸⁷ That is, the common ancestor of the Nuclear Cordilleran and Kalinga-Itneg languages.

reproduced here as (97-99).

Limos Kalinga

- (97) *singnad* Ø *dit nang-wa* Ø *tun pita?*
 IPW ABS NS PFV.AFOC-make ABS NS earth
 ‘Who made the earth? (lit.: ‘Who is that one out of sight that made the earth?’)’ (Ferreirinho 1993:28)
- (98) *ngan* Ø *dit maka-kan = taku?*
 IPW ABS NS ASSOC-eat-2PL[ABS]
 ‘Who will we eat with? (lit.: ‘Who is that one out of sight that we will eat with?’)’ (Ferreirinho 1993:52)
- (99) *ngadan* Ø *din bagbaga-on?*
 IPW ABS NS NMLZ~say-PFOC
 ‘What’s the discussion (about)? (lit.: What is that discussion there?)’
 (Ferreirinho 1993:104)

All the three interrogatives in (97-99) clearly contain a reflex of the word meaning ‘name’, which is *ngadan* in the variety described by Ferreira (1993). The initial *sing-* /siŋ-/ could be a combination of a personal nominal specifier *si* (Ferreirinho 1993:10), as in (100), followed by a “frozen enclitic ligature” (cf. Reid 2006 on Limos Kalinga nominal specifiers), similar to the one found on some of the common nominal specifiers, the proximate *tu-n* (97) and the distal (visible) *di-n* (99). However, given that the personal nominal specifier *si* “is never directly followed by a ligature elsewhere” in Kalinga languages, Lawrence Reid (p.c.) suggests that *sing-* is “more likely to be a reduced form of *sinu* > *sin* > *sing* / *ng*”.

Limos Kalinga (Ferreirinho 1993:11)

- (100) *kaysan* Ø *si* *Pedlo*
 left.AFOC ABS NS_{+PERSONAL} PROP
 ‘Pedro left.’

Obviously, it is the presence of the personal nominal specifier *si* that accounts for the human meaning ‘who?’ of the interrogative pronominal in (97). A similar development can be found in another Central Cordilleran language Amganad Ifugao, spoken immediately to the south of the Upper Tanudan Kalinga region. Thus, the Amganad Ifugao interrogative ‘who?’ is *hi ngadan* (Reid 1971:163), where *hi*, a regular reflex of **si*, must be a personal nominal specifier and *ngadan* is identical to the word meaning ‘name’.

4.2.2.1.3.1.4 Guinaang Kalinga

Guinaang Kalinga is a small Kalinga idiom spoken in *barrio* Guinaang a few

kilometers to the north from the town of Lubuagan. The Ethnologue classifies it as a dialect of Lubuagan Kalinga. Guinaang Kalinga uses one interrogative pronominal *sinu* both as ‘who?’ and ‘what?’ (Reid 1971:161, 163). The origins of this and similar ‘who?, what?’ interrogatives in South-Central Cordilleran languages have been discussed in Section III.4.2.2.1.2. In particular, it has been suggested that the lack of differentiation between ‘who?’ and ‘what?’ with such interrogatives is due to the peculiarities of the systems of nominal specifiers and case markers in the languages of Northern Luzon and the original semantics of the nominal specifier **si*. Table 2 and Table 3 in Section III.4.2.2.1.2 provide an overview of the systems of case markers and nominal specifiers of Guinaang Kalinga.

4.2.2.1.3.2 Nuclear Cordilleran languages

4.2.2.1.3.2.1 Lagawe and Bayninan Ifugao

According to the Ethnologue, Lagawe Ifugao is a variety of Tuwali Ifugao. It is spoken in Lagawe in the Southern Ifugao Province on Luzon Island somewhat to the south of the Kalinga region. Bayninan Ifugao is spoken in Bayninan, less than 20 kilometers to the north of Lagawe. Bayninan Ifugao appears to be classified by the Ethnologue as a dialect of a Southern Cordilleran language Keley-i Kallahan. However, as far as I can judge from the data found in Reid (1971:13), at least in its system of personal pronouns and demonstratives, Bayninan Ifugao looks much more similar to other Ifugao varieties than to Keley-i Kallahan.⁸⁸ Therefore, I consider it here as a Nuclear Cordilleran language.

According to Anne West (p.c.), Lagawe Ifugao uses only one interrogative pronominal *nganne* ‘who?, what?, which one? (person or thing)’ (101-102).

Lagawe Ifugao (Anne West, p.c.)

(101) *nganne* = *y nangat i he?a?*

IPW=[ABS]NS PFV.AFOC.bite OBL 2SG

‘Who/What bit you? (lit.: ‘Who/what is the one that bit at you?’)’

(102) *nganne* = *h tuwe = n kayiw?*

IPW=[ABS]NS this=LIG tree

‘What is this tree?’

Bayninan Ifugao has the same interrogative *nganne* as ‘who?, what?’, next to a dedicated ‘what?’ interrogative *nganu* (Reid 1971:161, 163, 233).

The Lagawe and Bayninan Ifugao interrogative *nganne* ‘who?, what?’ seems

⁸⁸ Lawrence Reid (p.c.) has confirmed that Bayninan Ifugao is probably misplaced in the Ethnologue and must be one of the Ifugao languages.

to reflect the same original structure as the Lower Tanudan Kalinga interrogative *ngai* ‘who?, what?’, discussed in Section III.4.2.2.1.3.1.2, i.e. [*ngadan* ‘name’ + a nominal specifier **di*]. In yet another Ifugao variety, Amganad Ifugao, the same form *nganne* means ‘what?’ (Reid 1971:161).

4.2.2.1.3.2.2 Northern Kankanay

Northern Kankanay is a relatively big Nuclear Cordilleran language spoken in an area to the south-west of Kalinga and to the west of Ifugao. According to Porter (1979:52) and Reid (1971:161, 163, 233), Northern Kankanay has two interrogative pronominals, *ngan* ‘what?’ and *sino* both ‘who?’ (103) and ‘what?’ (104).⁸⁹

Northern Kankanay (Porter 1979:26)

- (103) *sino* Ø *nan gawis ay maisolat is na?*
 IPW ABS NS be.good LIG be.written OBL this
 ‘What would be good to be written here?’
- (104) *sino* Ø *nan nangisolat is na?*
 IPW ABS NS wrote OBL this
 ‘Who wrote this?’

Apparently, the Northern Kankanay interrogative *ngan* ‘what?’ goes back to the word ‘name’ similarly to the Upper Tanudan Kalinga *ngan* ‘who?, what?’ and other similar interrogative pronominals of Central Cordilleran languages (cf. Sections III.4.2.2.1.1, III.4.2.2.1.3.1.1-III.4.2.2.1.3.1.3, III.4.2.2.1.3.2.1). The origins of South-Central Cordilleran ‘who?, what?’ interrogatives similar to the Northern Kankanay *sino* have been discussed in detail in Section III.4.2.2.1.2. In particular, it has been suggested that the lack of differentiation between ‘who?’ and ‘what?’ with such interrogatives is due to the peculiarities of the systems of nominal specifiers and case markers in the languages of Northern Luzon and the original semantics of the nominal specifier **si*. Table 6 and Table 7 provide an overview of the systems of case markers and nominal specifiers in Northern Kankanay.

⁸⁹ The data in Reid (1971) is from a Northern Kankanay variety spoken in Balugan. Porter (1979) herself does not further specify the dialect her data comes from, but according to Lawrence Reid (p.c.), Porter’s data are also from Balugan (locally known as Bugang), where she lived during her time with SIL in Northern Kankanay.

Table 6. Northern Kankanay case markers (based on Reid 2006)⁹⁰

TOP/PRED/ABS		GEN/ERG	OBL	LOC
Ø	after V:	= <i>n</i>	= <i>s</i>	= <i>d</i>
	after C:	Ø	<i>si</i> ~ <i>?is</i>	<i>?id</i>

Table 7. Northern Kankanay nominal specifiers (Reid 2006)

Indefinite	<i>di</i>
Definite ⁹¹	default <i>nan</i>
	recognitional <i>san</i>
	antedated <i>din</i>

As can be readily observed, Northern Kankanay has lost the nominal specifier **si*.⁹² The only noun phrase marker *si* that has remained in this language is the oblique case marker.

4.2.2.1.3.2.3 Balangao

Balangao is a small Nuclear Cordilleran language spoken in an area immediately to the south of the Lower Tanudan Kalinga speaking region. The Balangao dialect of Natonin uses one interrogative pronominal *heno* both as ‘who?’ and ‘what?’ (Reid 1971:161, 163).⁹³ The origins of South-Central Cordilleran ‘who?’,

⁹⁰ Reid’s (2006) Northern Kankanay data are “drawn primarily from” Hettick & Kent (1967) and Hettick & Wallace (1978), which describes the Northern Kankanay variety spoken in Balugan.

⁹¹ The term *recognitional* refers to an entity “within the recent common experience of speaker and addressee”, i.e. “the one that you and I have just been talking about, or experienced” (Reid 2006). Reid acknowledges Ruffolo (2005) and ultimately Himmelmann (1996) and Diessel (1999) as the sources for this term. The term *antedated* refers to “an entity that the speaker characterizes as being in the relatively distant past” (Reid 2006).

⁹² It should be mentioned that according to Lawrence Reid (p.c.), Kankanay languages never had “a reflex of **si* as an indefinite marker of genitive [i.e., ERG/GEN] NPs”, instead “the indefinite genitive form that is reconstructible to the parent language of the group is **di*.” However, this needs to be reconciled with the presence of *si*- on the interrogative *sino*. The first possibility would be that *si* became frozen on the interrogative in the period predating Proto Kankanay. Alternatively, *sino* could have been borrowed into Kankanay from some neighbouring languages.

⁹³ Note that this does not appear to hold true for all varieties of Balangao. Thus, the variety used in Shetler & Walrod’s (1983) paper on the structure of Balangao discourse seems to use

what?’ interrogatives similar to the Balangao *heno* have been discussed in detail in Section III.4.2.2.1.2. The explanation proposed appeals to the peculiarities of the systems of nominal specifiers and case markers in the languages of Northern Luzon and the original semantics of the nominal specifier **si*. Table 8 and Table 9 provide an overview of the systems of case markers and nominal specifiers in Balangao.

Table 8. Balangao case markers (based on Reid 2006)

TOP/PRED/ABS	GEN/ERG	OBL	LOC
\emptyset	after V: = <i>n</i>	= <i>h</i>	= <i>d</i>
	after C: \emptyset	? <i>ah</i>	? <i>ad</i>

Table 9. Balangao nominal specifiers (Reid 2006)

Indefinite	<i>di</i>
Definite	after V: = <i>n</i>
	after C: <i>hen</i>

According to Lawrence Reid (p.c.), Balangao *heno* is a regular reflex of **sinu*, with normal vowel-lowering in this context.

4.2.2.1.4 Southern Cordilleran languages

There are at least two Southern Cordilleran language having a ‘who?, what?’ interrogative, Kayapa Kallahan (proper) and Keley-i Kallahan.⁹⁴ The ‘who?, what?’ interrogative in Kayapa Kallahan has the form *hipa* (Reid 1971:161, 163), while in Keley-i Kallahan it is *hipa(?)* (Hohulin & Hale 1977; Reid 1971:161, 163).⁹⁵ The use of the Keley-i Kallahan interrogative is illustrated in

different interrogative pronominals.

⁹⁴ Kallahan is also known as Kalanguya. Kayapa Kallahan proper is spoken in Kayapa. The Keley-i Kallahan data is from the variety spoken in Caccajja Antipolo (Hohulin & Hale 1977:229). Note the spelling *Keley-i* with a lower case *-i* at the end. According to Lawrence Reid (p.c.), the spelling *Kelley-I* in the Ethnologue is not correct.

⁹⁵ A remark should be made on the Keley-i Kallahan form. The form of the interrogative in Hohulin & Hale (1977) is *hipa* while Reid (1971:161, 163) gives it as *hipa?*. This difference is probably accounted for by the fact that unlike Reid (1971), Hohulin & Hale (1977) do not mark glottal stops. Thus, indirect evidence for the final glottal stop in Hohulin & Hale’s (1977) form is provided by the gemination of the *d* of =*da* in (108). Another point to be made about the forms of the Keley-i Kallahan interrogative pronominal is that Reid (1971:161, 163) gives

(105-108).

Keley-i Kallahan

- (105) *hipa* Ø *hi* *ditten?*
 IPW ABS NS+PERSONAL DEM
 ‘Who is that (near you)?’ (Hohulin & Hale 1977:221)
- (106) *hipa* Ø *dimmen?*
 IPW ABS DEM
 ‘What is that ? (far from me and you)’ (Hohulin & Hale 1977:221)
- (107) *hipa*=*tu* *nanghelan* Ø *nunman?*
 IPW=ERG.3SG said ABS DEM
 ‘Why did he say that?’ (Hohulin & Hale 1977:225)
- (108) *hipadda* Ø *etan nangngel nem kinehing=da?*
 IPW=ERG.3PL ABS NS heard but disobeyed=ERG.3PL
 ‘Who were those that heard but disobeyed?’ (Hohulin & Hale 1977:214)

It is clear that the initial *hi-* here is a frozen regular reflex of the noun phrase marker **si*, while *-pa* is a reflex of one of the Proto Malayo-Polynesian ‘what?’ interrogatives **apa* (Blust *et al.* 2006).⁹⁶ The origins of this and similar ‘who?, what?’ interrogatives in South-Central Cordilleran languages have been discussed in detail in Section III.4.2.2.1.2. Thus, the lack of differentiation between ‘who?’ and ‘what?’ with such interrogatives has been explained by the peculiarities of the systems of nominal specifiers and case markers in the languages of Northern Luzon and the original semantics of the nominal specifier **si*, of which *hi* of *hipa(?)* is a regular reflex. Table 4 and Table 5 in Section III.4.2.2.1.2 provide an overview of the systems of case markers and nominal specifiers in Keley-i Kallahan. As has been pointed out in Section III.4.2.2.1.2, of particular interest is the fact that although in present-day Keley-i Kallahan *hi* is a personal nominal specifier, the interrogative *hipa(?)* containing a frozen nominal specifier *hi* means both ‘who?’ and ‘what?’, which suggests that earlier the nominal specifier at issue had a wider range of functions, not restricted to personal contexts.

hipa? as ‘what?’ and *hipa? humman* as ‘who?’, but this is not supported by the data in Hohulin & Hale (1977). Note in this respect that the Reid’s (1971:161, 163) form *hipa? humman* is literally just [IPW + that] (e.g., cf. Hohulin & Hale 1977:219) and similarly to (106) probably means ‘what is that (distant or out of sight)?’.

⁹⁶ Compare, for instance, the Indonesian interrogatives *si-apa* ‘who?’ and *apa* ‘what?’ (Austronesian, Malayo-Polynesian, Malayic; Blust *et al.* 2006).

4.2.2.2 Northern Cordilleran languages

The Northern Cordilleran languages of Northern Luzon are spoken in the north and north-east of Luzon Island. The internal classification of Northern Cordilleran languages as found in the Ethnologue is provided in Figure 8.

Figure 8. Northern Cordilleran languages according to the Ethnologue

Dumagat

- Northern
 - Paranan
 - Agta (Central Cagayan, Casiguran Dumagat, Dupaninan, Dicomay idioms)
 - Kasiguranin
- Southern: Agta (Camarines Norte, Umiray Dumagat, Alabat Island idioms)

Ibanagic

- Gaddang
- Ibanag
 - Atta (Pudtol, Pamplona, Faire idioms)
 - Villa Viciosa Agta
 - Ibanag
 - Itawit
 - Yogad
- Isnag
 - Isnag
 - Adasen Itneg
- Moyadan Itneg

There is at least one Northern Cordilleran idiom, Dibagat Isnag, that appears to have an interrogative pronominal used indiscriminately as ‘who?’ or ‘what?’. Dibagat Isnag is a variety of Isnag (or Isneg) spoken in Dibagat (Kabugao municipality of the province of Apayao) in the north of Luzon Island. According to G. Richard Roe’s data cited in Reid (1971:161, 163, 233, 235), the interrogative pronominals of Dibagat Isnag are *na:gan* ‘what?’, *ʔiʔin* ‘who?’ and *ʔin* ‘who?, what?’.

The only source with examples of the Isnag interrogatives I was able to consult is Barlaan’s (1977) paper on Isnag discourse. However, given that Barlaan does not indicate the exact variety of Isnag he is describing, this paper may pertain to a dialect other than Dibagat Isnag. Still, it may be illustrative to consider the examples of interrogative pronominals that can be found in Barlaan’s paper. Therefore, in what follows, I will first present the examples with interrogative pronominals found in Barlaan (1977) and then proceed to a discussion of the possible origins of the ‘who?, what?’ interrogative *ʔin*.

In Barlaan (1977), I have found examples with *iin* as ‘who?’ (109), *in* as ‘what?’ (110), *an* as ‘what?’ (111).⁹⁷

Isnag (Austronesian, Malayo-Polynesian, Northern Cordilleran; Philippines)

(109) *iin = ka?*

IPW=2SG[ABS]

‘Who are you?’ (Barlaan 1977:129)

(110) *in Ø da kagkagiyan ne Kikit tin?*

IPW ABS NS NMLZ~say.PFOC GEN.NS PROP DEM

‘What is Little Finger saying now? (lit.: What are the words of Little Finger here?)’ (Barlaan 1977:130)

(111) *an Ø da kagkagiyan ne Kikit ti?*

IPW ABS NS NMLZ~say.PFOC GEN.NS PROP DEM

‘What is Little Finger saying? (lit.: What are the words of Little Finger here?)’ (Barlaan 1977:131)

Besides, there is one confusing example (112) with a form *inna*, glossed as ‘who?’ but translated as ‘what?’. In (112), I reproduce the original glossing and provide a possible alternative.

Isnag (Austronesian, Malayo-Polynesian, Northern Cordilleran; Philippines)

(112) *oy in = na na:d da kuwa ya:n ne?*

hey who UNCERTAINTY LINK HESITATION this

hey IPW=3SG[ABS] UNCERTAINTY [ABS]NS HESITATION this GEN.NS

‘[There is there a Wood Borer. “Gnaw, gnaw,” the Wood Borer said (as he) went around the edge of Gisorab’s (house).] “Hay I wonder what that (is),” [he said. “Who are you?” “I am a Wood Borer.”]’ (Barlaan 1977:128)

In all probability, *inna* is in fact the interrogative *in* followed by an encliticized third person singular “subject/genitive” pronoun *na*.⁹⁸ The final *ne* looks like a genitive-ergative personal nominal specifier *ne*, as in (113), but used as a

⁹⁷ Note that Barlaan (1977) does not mark glottal stops, so that in fact his *iin* is very likely to be /ʔiʔin/, and his *in* /ʔin/ and his *an* /ʔan/.

⁹⁸ In other Northern Luzon languages “subject/genitive” would normally imply ergative-genitive (cf. Reid & Liao 2004; Reid 2006). However, in the texts in Barlaan (1977) the “subject/genitive” *na* sometimes also appears to function as absolutive. Note, though, that in the latter function the third person singular pronoun seems to be more commonly left unexpressed, particularly in verbal clauses. It should also be mentioned that according to Lawrence Reid (p.c.), =*na* of *in = na* is more likely to be a demonstrative *na* cognate to the Bontok proximal demonstrative *na*, even though “the Isnag semantic equivalent is *ya* ‘this’”.

pronominal.

Isnag (Austronesian, Malayo-Polynesian, Northern Cordilleran; Philippines)

(113) *ta:n ya:n ne Akakutan?*

there.is this GEN.NS+PERSONAL PROP

‘There is there a Wood Borer (lit.: ‘There is this of Wood Borer)’ (Barlaan 1977:128)

The element *da* may be a plural nominal specifier followed by a hesitation particle *kuwa*, which is often translated as ‘what-you-may-call-it?’ (cf. also Reid 1971:161). Thus, the whole sentence should probably be translated as something like ‘Hey, who/what could it (be)? This what-you-may-call-it, this one (person)?’. As such, the context of (112) does not provide enough information to decide whether *in* here means ‘what?’ or ‘who?’. However, if *ne* is indeed a pronominally used personal marker, the ‘who?’ interpretation of *in* in (112) would be more natural, agreeing thus with the gloss ‘who?’ provided by Barlaan himself. Taken together with the use of *in* as ‘what?’ in (110), this would suggest that *in* is a ‘who?, what?’ interrogative in the Isnag variety described in Barlaan (1977).

Finally, besides the examples presented above, I have encountered two examples with an omitted interrogative pronominal (114a) and (114b).

Isnag (Austronesian, Malayo-Polynesian, Northern Cordilleran; Philippines)

(114) a. *Ø da kagkagiyan mu Kikit?*

ABS NS NMLZ~say.PFOC GEN.2SG PROP

‘What are you saying Little Finger? (lit.: ‘These words of yours Little Finger (are what)?..’)

b. *kagkagiyan ne Kikit ti?*

NMLZ~say.PFOC GEN.NS PROP DEM

‘What is Little Finger saying? (lit.: ‘Little Finger is saying here (what)?..’)

In (114a), the element *da* glossed by Barlaan as ‘what?’ must be a plural nominal specifier. In this respect, compare (114a) to (110) and (111).

Let us now discuss possible origins of the interrogative pronominal *ʔin* ‘who?, what?’. Two more or less equally plausible hypotheses can be advanced here. First, *ʔin* may go back to the Proto Austronesian interrogative **i-nu* ‘where?’ (cf. Blust *et al.* 2006), which in Isnag must have first developed into a selective interrogative ‘which one? (person or thing)’ and later has been extended to non-selective contexts as ‘who?’ or ‘what?’. Subsequently, the interrogative *ʔi-ʔin* ‘who?’ has been derived from *ʔin* ‘who?, what?’ by means of an earlier personal marker *ʔi-*, which is not productive in Isnag anymore (see below).

This hypothesis looks all the more plausible because the use of ‘where?’ as

‘which one?’ is not uncommon in the Philippine Austronesian languages (a further extension to non-selective contexts is also semantically easy to conceive). Thus, in Lower Tanudan Kalinga *dinu* is used both as ‘where?’ and ‘which one?’, although more commonly as the latter (Central Cordilleran; Glenn & Jewell Machlan, p.c.).⁹⁹ In Karao (Southern Cordilleran; Sherri Brainard, p.c.), the only way to ask ‘which one?’ is to use the interrogative *to* ‘where?’. The use of ‘where?’ as ‘which one?’ is also very common among the Austronesian languages of Taiwan (cf. Huang *et al.* 1999:654, 663-664), as in Tsou *nenu* ‘where?, which one?’ (Austronesian, Tsouic; Taiwan) or Paiwan *inu* ‘where?, which one?’ (Austronesian, Paiwan; Taiwan). Numerous further examples from all over Austronesian can be added here.

An alternative hypothesis would be that there was no preceding locative stage. That is, the interrogative **ʔi-nu*, of which *ʔin* is the reflex, was not a locative interrogative ‘where?’ but a ‘who?, what?’ interrogative structurally identical to the *si-nu* (and the like) ‘who?, what?’ interrogatives in South-Central Cordilleran languages discussed in Section III.4.2.2.1.2. That the ‘where?’ hypothesis discussed above is not the only possibility has been pointed out to me by Lawrence Reid (p.c.). Thus, he notes that “in the parent of the Philippine languages **ʔi* is reconstructible with various functions”, locative preposition being only one of them (for more details cf. Reid 1979:11-16; Reid & Liao 2004:469). What is more, **ʔi* is also reconstructible as both a personal and a common noun phrase marker and it must have “preceded the use of **si* as personal noun marker”. In other words, it appears that **ʔi* is largely an earlier functional equivalent of **si*, so that **ʔi-nu* ‘who?, what?’ could be accounted for in the same way as *si-nu* (and the like) ‘who?, what?’ interrogatives in South-Central Cordilleran languages discussed in Section III.4.2.2.1.2.

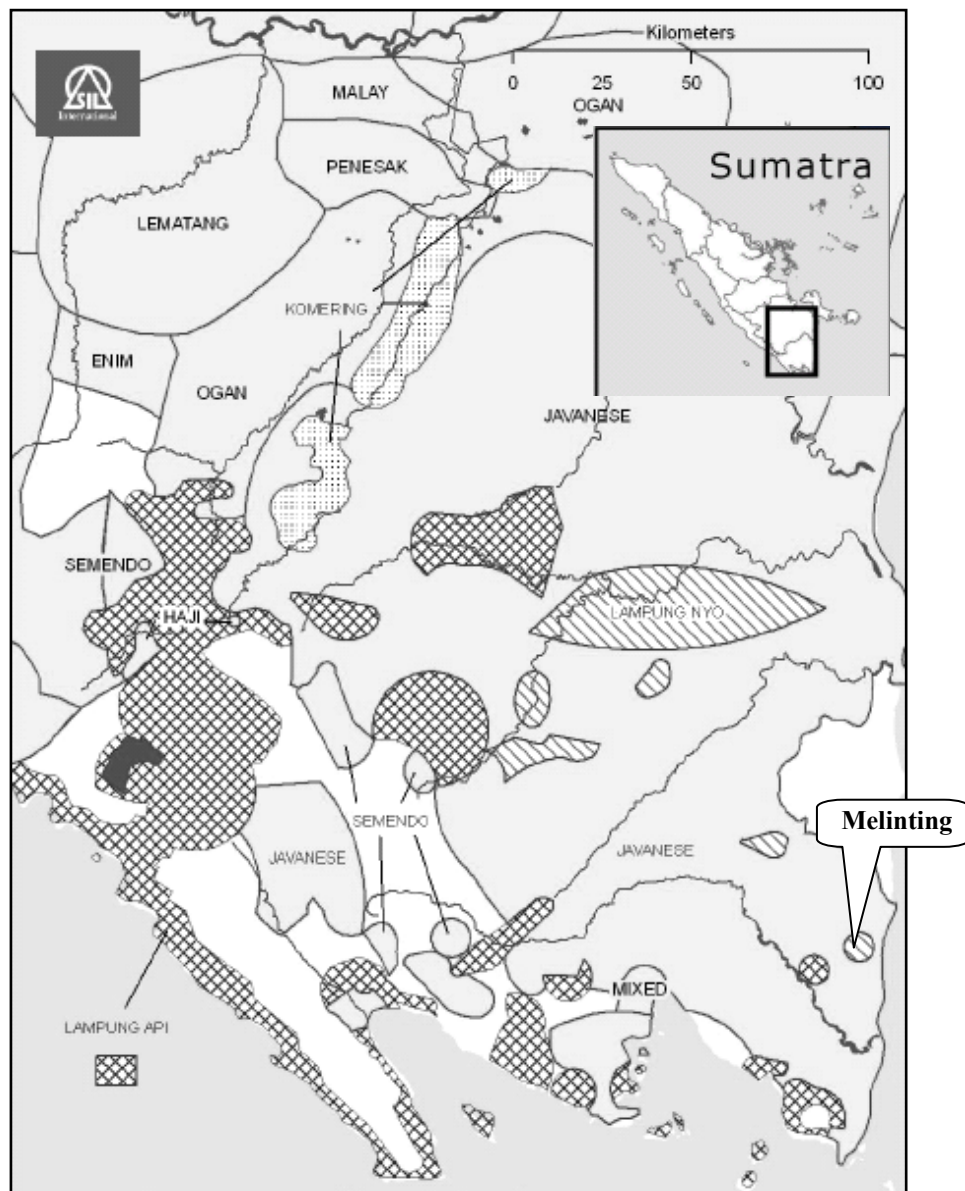
4.2.3 Lampungic languages

Lampungic languages are spoken by some one and a half million people in the south of Sumatra Island (Indonesia). According to the Ethnologue, they form an isolate subgroup within the Malayo-Polynesian branch of the Austronesian family. However, Anderbeck (2006) suggests that Lampungic should probably be

⁹⁹ A more common form for ‘where?’, *dimman*, must result from a fusion of *dinu* ‘where?’ with the following distal demonstrative *man*. The Lower Tanudan Kalinga interrogative *dinu* itself can be compared to the interrogative *di ʔano* ‘where?’ (lit. ‘at what?’) in Binongan Itneg, a closely related Central Cordilleran language. The element *di* is a locative case marker, widespread in Northern Luzon languages (Reid 2006:42), and *ʔano* is a reflex of the Proto Malayo-Polynesian **-anu* ‘what?’ (Blust *et al.* 2006), which is also common in the languages of the Philippines.

grouped together with Sundanese. The internal classification of Lampungic has been the object of Hanawalt's (2006) paper, where several possibilities have been considered. Hanawalt *et al.* (in prep., cited via Anderbeck 2006:8-9) favour the option dividing Lampungic into three major dialect clusters, Lampung Api (or Pesisir), Komering (or Komering-Kayu Agung) and Lampung Nyo (or Abung). The first two clusters, Api and Komering, do not show "any strong linguistic differences [...] their relationship is more of a language chain than two completely separate clusters", while the third one, Nyo, is quite different (Anderbeck 2006:9). The distribution of the three major Lampungic dialect clusters can be found on Map 8.

Map 8. Distribution of the major subgroups within Lampungic (based on Anderbeck 2006:9, 13; Hanawalt 2006:27)



There is one Lampungic idiom with an interrogative pronominal used indiscriminately as ‘who?’ or ‘what?’. This idiom is spoken in Nibung village and constitutes the Melinting variety of the Lampung Nyo dialect cluster. The ‘who?, what?’ interrogative of Melinting has the form *apəɔ* (Anderbeck 2006:71). The lack of differentiation between ‘who?’ and ‘what?’ with this interrogative can be explained in two ways. First, it may be due to the loss of a frozen reflex of the preposed “personal” marker **si* that used to distinguish the interrogative ‘who?’ from the interrogative ‘what?’. Second, it may be due to the fact that the Melinting form *apəɔ* originates in a construction of the structure [‘what?’ + a demonstrative or personal pronominal], literally meaning something like ‘what one?’, i.e. ‘which one? (person or thing)’, which later has been extended to non-selective contexts as both ‘who?’ and ‘what?’.

Let us first consider the hypothesis explaining the lack of differentiation between ‘who?’ and ‘what?’ with the Melinting interrogative *apəɔ* through the loss of a frozen reflex of the preposed “personal” marker **si* that used to distinguish the interrogative ‘who?’ from the interrogative ‘what?’. The respective interrogative ‘what?’ is a reflex of the Proto Malayo-Polynesian **apa* ‘what?’ (cf. Blust *et al.* 2006). That is, in an earlier stage ‘who?’ and ‘what?’ in the variety of Lampungic at issue may have been differentiated in the same way as in Indonesian, which has *si-apa* ‘who?’ vs. *apa* ‘what?’ (Austronesian, Malayo-Polynesian, Malayic; Blust *et al.* 2006). The further loss of *si-* in Melinting could have proceeded as follows. First, the vowel *i* was dropped resulting in **sapa*, similarly to most other Lampung varieties (Table 10 below) or the neighbouring Javanese *sapa* ‘who?’ vs. Indonesian *siapa* ‘who?’.¹⁰⁰ Then, the initial *s-* became *h-*, a change occurring sporadically in Lampungic and more frequently in the neighbouring languages.¹⁰¹ In turn, the initial *h-* was deleted, which is a regular phonological process in Lampung Nyo (see Anderbeck 2006:20). At a certain point, the final *-a* of the interrogative became *o* and later became diphthongized to *əɔ*, which is a regular phonological change in Melinting and some other Lampung Nyo dialects (see Anderbeck 2006:22-24). In other words, the development of the ‘who?, what?’ interrogative in Melinting may be due to a concurrence of some more or less regular phonological developments. In this respect, compare a perhaps similar development in the Melinting demonstrative pronominal *inəɔ* ‘that’, which can be compared to Komering Adumanis *sina*, Komering Ilir *sino*, Lampung Api Belalau *hin:o*, and Lampung Api Jabung *inə*.

Let us now consider the hypothesis explaining the lack of differentiation

¹⁰⁰ The interrogative ‘what?’ is *apa* both in Javanese and Indonesian.

¹⁰¹ Cf. Anderbeck (2006:24): “**s* is a remarkably stable phoneme when one considers all the permutations which occur in Sumatran Malay. It is rarely elided or weakened to *h*”.

between ‘who?’ and ‘what?’ with the Melinting interrogative *apəɔ* by bringing it back to an interrogative constructed as [‘what?’ + a demonstrative or personal pronominal] and literally meaning ‘what one?’, i.e. ‘which one? (person or thing)’. To begin with, let us consider the interrogative pronominals ‘who?’ and ‘what?’ from various Lampungic idioms summarized in Table 10.

Table 10. Lampungic interrogative pronominals ‘who?’ and ‘what?’ (Anderbeck 2006:71)

	‘what?’	‘who?’
Lampung Api		
Ranau	<i>ʒpi</i>	<i>sʒpɜ</i>
Jabung	<i>api</i>	<i>sapɔ</i>
all other dialects	<i>api</i>	<i>sapa</i>
Komerling		
Iilir	<i>apiya</i>	<i>sapa</i>
Kayu Agung Asli	<i>ʌpi</i>	<i>sape</i>
Kayu Agung Pendatang	<i>oni</i>	<i>sape</i>
all other dialects	<i>api</i>	<i>sapa</i>
Lampung Nyo		
Kotabumi	<i>ɲɔ</i>	<i>siapɔ</i>
Menggala	<i>ɲow</i>	<i>sapow</i>
Sukadana	<i>ɲəw</i>	<i>apəɔ</i>
Melinting	<i>apəɔ</i>	

On the basis of these forms, Anderbeck (2006:71) reconstructs three interrogative pronominals for Proto Lampungic, viz. **api* ‘what?’, **əɲa* ‘what?’ and **sa-apa* ‘who?’, where **sa-*, together with another reconstructed form **h-*, is said to be “some kind of personal marker” (2006:83). Reflexes of the two “personal” markers can also be found on some personal pronouns and on demonstrative pronominals (Anderbeck 2006:83).

I believe that Anderbeck’s reconstructions can be improved on several points. First, as suggested by the Kotabumi form *siapɔ* ‘who?’, this and similar interrogatives should rather be reconstructed with the initial “personal” marker **si-* and not **sa-*. Further support comes from the forms of the Lampungic demonstrative and personal pronominals containing a reflex of the “personal” marker. All these forms have *i* in the first syllable and not *a* (cf. Anderbeck 2006:69-70, 83). It seems reasonable to conclude that the “personal” marker should be reconstructed as **si*, with a later lenition to **hi*, loss of the vowel *-i*

and/or *h*.¹⁰² Morphemes of similar forms are found in other Malayo-Polynesian languages, cf. Section III.4.2.2.1.2 for more details on the “personal” marker **si*.

The second point is about the form **api* ‘what?’ reconstructed by Anderbeck. As suggested by the Komering Ilir form *apiya* ‘what?’, this and similar interrogatives represent reflexes of the Proto Malayo-Polynesian **apa* ‘what?’, also found in various Lampungic *sapa* (and the like) ‘who?’ interrogatives, followed by some other element. The latter element was in all probability a demonstrative. In modern Lampungic idioms, it has been preserved as the third person singular pronoun, its most common form being *i(y)a* and the like, with a regular rounding and subsequent diphthongization of the final vowel in some varieties, such as Melinting *iəɔ* (cf. Anderbeck 2006:70).¹⁰³ That is, instead of **api* ‘what?’ we should reconstruct the combination **apa-ia* (or the like), literally meaning something like ‘what one?’ (or ‘its what?’).¹⁰⁴ Such a compositional origin of the interrogative *api* (and the like) ‘what?’ in Lampungic also squares well with the fact that the *api*-like interrogatives are exceedingly rare in the Malayo-Polynesian languages, while the *apa*-like interrogatives are so frequent that they allow for **apa* ‘what?’ to be reconstructed for Proto Malayo-Polynesian (cf. Blust *et al.* 2006).

Melinting *apəɔ* ‘who?, what?’ and Sukadana *apəɔ* ‘who?’ may go back either to **apa* ‘what?’ or **apa-ia* ‘what one?’ (i.e., ‘which one?’), because both developments are phonologically plausible. If they go back to **apa* ‘what?’ then the only way to explain their shift/expansion to ‘who?’ in Sukadana/Melinting would be to assume a loss of the “personal” marker **si* in the overtly marked interrogative **si-apa* ‘who?’ > **apa* (cf. the discussion above), which then was

¹⁰² I suspect that the form of Anderbeck’s reconstruction of the marker in question as **sa-* has been influenced by the existence of a demonstrative *sa* ‘this’ in some varieties of Lampung (Anderbeck 2006:69).

¹⁰³ Anderbeck (2006:70) compares the Lampungic pronominal to the Proto Malayo-Polynesian pronoun **si-ia*. Note, however, that **ia* itself, or rather **ya* is also reconstructed as a distal demonstrative in the Northern Luzon Philippine languages (cf. Reid 2006:26). Given that at the same time no third person singular pronominal in Anderbeck’s Lampungic data begins with an *s-*, it seems reasonable to assume that in Lampungic this third person pronominal goes back directly to a demonstrative **ya* (or the like).

¹⁰⁴ Interrogative pronominals based on the same or similar structure, [‘what?’ + a demonstrative or personal pronominal], can be found in other Malayo-Polynesian languages as well. Thus, Blake (1906:362-363) suggests that Pangasinan *an-to* ‘what?’ literally means ‘its what?’, where *an-* ‘what?’, is followed by a genitive third person pronoun *to* (Northern Philippine, Northern Luzon, Southern Cordilleran). The same form *to* also makes part of some of the Pangasinan demonstratives, and in all probability is a reflex of a proximate demonstrative **tu* (cf. Reid 2006:20).

no more formally differentiated from **apa* ‘what?’. Otherwise, if Melinting *apəp* ‘who?, what?’ and Sukadana *apəp* ‘who?’ go back to **apa-ia* ‘what one?’ (i.e., ‘which one?’), just like *api*-like interrogatives in the other Lampungic varieties, the differences in the meanings of the reflexes of **apa-ia* in Melinting, Sukadana and the other Lampungic varieties can be explained as follows. In most Lampungic varieties, **apa-ia* was understood more literally as ‘*what* one?’ and was therefore confined to questions about things. In Melinting, **apa-ia* was understood as broadly selective ‘which one? (person or thing)’. Sukadana might have been similar to Melinting before but because of the existence of a dedicated ‘what?’ interrogative *ɲəw* (as in Lampung Nyo Kotabumi and Menggala), the Sukadana reflex of **apa-ia* was confined to the human meaning ‘who?’.

The third and last remark about Anderbeck’s reconstruction concerns the form **əɲa* ‘what?’ reconstructed by Anderbeck. Like *api*, the interrogative pronominals of such a form are exceedingly rare in Malayo-Polynesian. I believe that the forms like Lampung Nyo Kotabumi *ɲə* ‘what?’ reflect a similar original structure as the *api*-like interrogatives, but are based on a reflex of another Proto Malayo-Polynesian ‘what?’ interrogative root **-anu* (cf. Blust *et al.* 2006). That is, *ɲə* and similar interrogatives should be reconstructed as a combination of **anu* (or the like), which is a reflex of the Proto Malayo-Polynesian **-anu* ‘what?’ and the demonstrative **ia* (or the like; see above, in particular footnote 103). The final vowel of **anu* was later dropped and **n* was palatalized into *ɲ* before **ia*. This hypothesis has two advantages. First, through bringing the *api*- and *ɲə*-like interrogatives to a similar original structure, it makes the system of the Lampungic interrogative pronominals more internally consistent and uniform.¹⁰⁵ Second, it can help us to account for the Komering Kayu Agung Pendatang form *ɔɲi* ‘what?’ in a straightforward way. The final *i* of *ɔɲi* will then be explained in the same way as the final *i* of *api*-like interrogatives in the other Komering dialects.

By way of conclusion, there are arguments to support both hypotheses

¹⁰⁵ As can be observed in Blust *et al.* (2006), the overall geographical distribution of the two ‘what?’ interrogatives reconstructed for Proto Malayo-Polynesian, **apa* and **-anu*, suggests that **apa*-based interrogative pronominals in Lampungic may be later loans from neighbouring languages, such as Malay, Ogan, etc. Thus, **anu*-based ‘what?’ interrogatives are most common in the Austronesian languages of the Philippines and are also found in Taiwan, as well as here and there in other parts of the Austronesian world, while **apa*-based ‘what?’ interrogatives are most common in the Central-Eastern Malayo-Polynesian languages, spoken in Indonesia and Oceania. The older status of **anu*-based ‘what?’ interrogatives in Lampungic follows then from the fact that Taiwan is commonly viewed as the Proto Austronesian homeland, with the Philippines being the first stop on a way to the Austronesian expansion into Indonesia, Oceania and Madagascar.

proposed above to account for the lack of differentiation between ‘who?’ and ‘what?’ with the Lampung Nyo Melinting interrogative *apəə*. However, all things being equal, I am more inclined to follow the hypothesis bringing *apəə* back to a construction of the structure [‘what?’ + a demonstrative or personal pronominal], because it involves more regular phonological developments than the hypothesis suggesting the loss of the initial **si-* and it would square better with the existence of similar evolutions in some other Austronesian languages (cf. Sections III.4.2.4-III.4.2.6).

4.2.4 Choiseul languages

The Choiseul languages (Babatana, Ririo, Vaghua and Varisi) are spoken on Choiseul Island (Solomon Islands). According to the Ethnologue, they form a lower-level subgrouping of the Meso Melanesian languages of the Western Oceanic group of the Eastern Malayo-Polynesian branch of the Austronesian family. There is at least one Choiseul idiom that may have an interrogative pronominal used indiscriminately as ‘who?’ or ‘what?’. According to Blust *et al.* (2006), Sisingga, classified by the Ethnologue as a dialect of Babatana, uses one interrogative *anda* for both ‘who?’ and ‘what?’. The Sisingga data are said to represent the variety spoken in the village of Gagara.

The interrogative *anda* appears to go back to a combination of a reflex of the Proto Malayo-Polynesian **-anu* ‘what?’ (Blust *et al.* 2006) followed by a demonstrative *ta* ‘this (here)’ (Blust *et al.* 2006). Literally, the original construction must have meant something like ‘what one?’, i.e. ‘which one? (person or thing)’. Alternatively, if the demonstrative was adverbial and not pronominal, the construction is more likely to have originally meant ‘where?’, with a later shift from ‘where?’ to ‘which one?’.¹⁰⁶ Subsequently, ‘which one?’ has been extended to non-selective contexts both as ‘who?’ and ‘what?’.

Interrogative pronominals based on a similar structure, i.e. [‘what?’ + a demonstrative or personal pronominal], can be found in other Malayo-Polynesian languages as well. For instance, Blake (1906:362-363) suggests that Pangasinan *an-to* ‘what?’ literally means ‘its what?’, where *an-* ‘what?’, which looks like a reflex of Proto Malayo-Polynesian **-anu* ‘what?’,¹⁰⁷ is followed by a genitive

¹⁰⁶ The use of ‘where?’ as ‘which one?’ is rather common both cross-linguistically and within Austronesian (cf. Section III.4.2.2.2 for some examples from Austronesian languages).

¹⁰⁷ It should be mentioned, however, that according to Reid (1979:8, 10) the initial *an-* in Pangasinan *an-to* is not a reflex of **anu* but goes back to **ŋajan* ‘name, what?’ (see Section III.4.2.2.1.1 on similar interrogatives in other Philippine languages), so that *an-to* is cognate to such interrogatives as *ŋan-to* ‘what?’ in Ibaloi, another Southern Cordilleran language. Although *an-* of Pangasinan *an-to* may indeed be a reflex of **ŋajan* rather than **anu*, the same

third person pronoun *to* (Northern Philippine, Northern Luzon, Southern Cordilleran). The same form *to* also makes part of some of the Pangasinan demonstratives, and in all probability is a reflex of a proximate demonstrative **tu* (cf. Reid 2006:20).¹⁰⁸ Consider also the Lampungic interrogative pronominals *api(ya)* and *ɲɔ*, discussed in Section III.4.2.3, which in all probability also go back to the structure [‘what?’ + a demonstrative or personal pronominal].

That the original meaning of the Sisingga *anda* may have also been ‘where?’ instead of ‘what one?’ is suggested by the fact that for instance in many Philippine languages similar interrogatives often mean ‘where?’, as e.g. Kapampangan *anta* ‘where?’ (Blust *et al.* 2006), and sometimes both ‘where?’ and secondarily ‘what?’, as Maranao *anda* (Malayo-Polynesian, Southern Philippine; Blust *et al.* 2006) or Bolaang Mongondow *onda* (Malayo-Polynesian, Sulawesi; Blust *et al.* 2006). Maranao and Bolaang Mongondow also have dedicated ‘what?’ interrogatives, *ai* and *onu* respectively. Numerous further examples can be found among the Austronesian interrogative forms collected in Blust *et al.* (2006).

4.2.5 *South Huon Gulf languages*¹⁰⁹

South Huon Gulf languages are spoken in the Morobe Province of Papua New Guinea on the southern coast of Huon Gulf. According to the Ethnologue, they form a lower-level subgrouping of the North New Guinea languages of the Western Oceanic group of the Eastern Malayo-Polynesian branch of the Austronesian family. There is at least one South Huon Gulf language that appears to have an interrogative pronominal used indiscriminately as ‘who?’ or ‘what?’. According to Bradshaw (2001) and Joel Bradshaw (p.c.), Iwal, a South Huon Gulf language spoken by a couple of thousand people, has an interrogative pronominal *ret* “that can be translated ‘what’ as well as ‘who’, depending on the context”, as in (115) and (116). Furthermore, *ret* can be used attributively in phrases as *amol ret* ‘what person?, who?’ or *gen ret* ‘what thing?, what?’.

Iwal (Austronesian, Eastern Malayo-Polynesian, Western Oceanic, North

hypothesis appears to be hardly tenable for *an-* of Sisingga *an-da*. Unlike with Pangasinan, there appear to be no languages within a reasonable distance from Sisingga to support this alternative hypothesis. Anyway, in the case of Sisingga it is not so important after all, whether *an-* goes back to **anu* or **ɲajan*, what matters is that in either case *an-* would go back to an earlier interrogative ‘what?’.

¹⁰⁸ Interestingly, Reid (2006:38) has also reconstructed a Proto Northern Luzon “near addressee” demonstrative **ta*, identical to the Sisingga *ta* ‘this (here)’.

¹⁰⁹ I would like to thank Joel Bradshaw for sharing with me his data on Iwal and several other Huon Gulf languages, helpful comments and interesting discussion.

New Guinea, South Huon Gulf; Papua New Guinea)

(115) *ret g-i-nei beti yem u-pelk?*

IPW PST-3SG-say so 2PL 2-fear

‘Who warned you to flee?’ (Matthew 3:7 via Joel Bradshaw, p.c.)

(116) *yem u-pelk dang etok ve ret ane?*

2PL 2-fear like that for IPW of

‘Why (i.e. ‘for what reason’) are you so afraid?’ (Matthew 8:26 via Joel Bradshaw, p.c.)

Besides *ret*, Iwal has a special plural human interrogative pronominal *asav* ‘who (PL)?’.

The exact origins of the Iwal interrogative *ret* are not immediately traceable because the interrogatives of exactly this form appear to be exceedingly rare in Austronesian. Thus, the closest parallel I found in Blust *et al.* (2006) is the Koronadal Bilaan interrogative *det* ‘what?’ (Malayo-Polynesian, South Mindanao; Philippines). However, there are several sporadic interrogative pronominals ending in *-re* and *-de* mostly among the Central-Eastern Malayo-Polynesian languages, e.g. Alune *sare* ‘what?’ and *side/sire* ‘who?’ (Central Malayo-Polynesian, Central Maluku; Indonesia; Blust *et al.* 2006), Bileki Nakanai *e-re* ‘who?’ vs. *l-ava* ‘what?’ (Eastern Malayo-Polynesian, Western Oceanic, Meso Melanesian; Papua New Guinea; Blust *et al.* 2006), Babatana *ande* ‘who?’ vs. *ava* ‘what?’ (Eastern Malayo-Polynesian, Western Oceanic, Meso Melanesian; Solomon Islands; Blust *et al.* 2006), Amba (or Nembao) *vare* ‘what?’ vs. *nía* ‘who?’ (Eastern Malayo-Polynesian, Remote Oceanic; Solomon Islands; Blust *et al.* 2006).

Two things strike the eye in these forms. First, *an-* in Babatana *ande* ‘who?’ and *va-* in Amba *vare* ‘what?’, Babatana *ava* ‘what?’, *-ava* in Bileki Nakanai *l-ava* ‘what?’ are most likely reflexes of the two Proto Malayo-Polynesian ‘what?’ interrogatives **-anu* and **apa* respectively. In turn, the final elements *-de(t)*, *-re(t)* resemble some common demonstrative roots or combinations thereof. For instance, various Philippine Northern Luzon languages have distal nominal specifiers of the forms *dí*, *ta* and *dít(a)*, the first going back to a distal demonstrative, the second to a medial/“near speaker” demonstrative and the third to a combination of the two (cf. Reid 2006).¹¹⁰ If we add to this that

¹¹⁰ Forms like *-re* or *-de* without a final consonant may also represent a combination of different demonstrative elements, comparable for instance to **ta* + **ya*, a medial/“near addressee” and a distal demonstrative root respectively in Northern Luzon languages (cf. Reid 2006). In this respect, consider also such interrogatives as Hanunoo *híntay* ‘what?’ and *síntay* ‘who?’ (Meso Philippine; Blust *et al.* 2006), Kiribati (Central-Eastern Oceanic; Kiribati, Solomon Islands & Fiji; Blust *et al.* 2006) or Maranao (Southern Philippine; Blust *et al.* 2006) *antai* ‘what?’.

interrogative pronominals based on the structure [‘what?’ + a demonstrative or personal pronominal] can be found in several Malayo-Polynesian languages,¹¹¹ a conclusion that suggests itself is that in all probability, both the Koronadal Bilaan interrogative *det* and the Iwal interrogative *ret* go back to the same structure [*-*anu* or **apa* ‘what?’ + **di-ta*, a fused combination of two non-proximate demonstrative roots],¹¹² which literally meant something like ‘what one?’, i.e. ‘which one?’. The original interrogative root has later disappeared, most likely following an evolution similar to that of Italian *cosa?* ‘what?’ that has developed from *che cosa?* ‘what thing?’. The only difference between the Koronadal Bilaan and Iwal interrogatives is that apparently, in Iwal it was originally possible to use the structure ‘what one?’ in selective questions both about persons and things, while in Koronadal Bilaan ‘what one?’ was understood more literally and was therefore confined to questions about things.

The hypothesis proposed above is appealing for several reasons. First of all, such a compositional origin of the interrogatives *det/ret* (and the like) squares well with the fact that the interrogatives of this and similar forms are very rare in the Malayo-Polynesian languages and occur in languages that are both geographically and genetically distant. This hypothesis implies just parallel developments and thus exempts us from the necessity to reconstruct such an interrogative for Proto Malayo-Polynesian. Second, as mentioned above interrogatives based on the same original structure appear to exist in other geographically and genetically distant Austronesian languages. In this respect, it is also illustrative to mention that for instance, Sarangani Bilaan, closely related to Koronadal Bilaan (cf. above in the present section), has an absolutely different interrogative ‘what?’, *tan* instead of *det*, which however again strikingly resembles demonstratives in some other Philippine languages, such as Karao *tan* ‘this (near you)’ (Northern Philippine; Sherri Brainard, p.c.).¹¹³ Finally, the development of **di-ta* to *ret* in Iwal is also plausible phonologically. Compare in this respect the reflexes of the Proto Austronesian **t-ina* ‘mother’ in some of the Huon Gulf languages: Numbami *tina*, Yabem *têna*, but Wampar *rena-n* (Blust *et al.* 2006).

¹¹¹ Consider the discussion of the Lampungic interrogative pronominals *api(ya)* and *ɲɔ* in Section III.4.2.3 or of the Sisingga interrogative *anda* ‘who?, what?’ discussed in Section III.4.2.4.

¹¹² In Northern Luzon languages reflexes of **dí* are usually distal demonstratives, while **ta* is reconstructed as a medial or “near addressee” demonstrative (cf. Reid 2006).

¹¹³ See Reid (2006) for other examples and for the reconstruction of *tan* as a combination of the medial/“near addressee” demonstratives **ta* and “a frozen final ligature” (cf. Reid 2006).

4.2.6 Tsouic languages¹¹⁴

The Tsouic languages (Tsou, Kanakanabu and Saaroa) are spoken in a compact mountainous region in south-central Taiwan, cf. Map 9. The received view is that Tsouic languages constitute a branch of Austronesian.¹¹⁵ Tsouic languages are subdivided in two groups: (i) Kanakanabu and Saaroa, two moribund languages spoken only by a handful of old people, and (ii) Tsou, counting probably around two thousand speakers. Kanakanabu and Saaroa are quite different from Tsou. In part, this is probably due to the fact that Kanakanabu and Saaroa have been strongly influenced by Siraya (Austronesian, East Formosan), Rukai (Austronesian, Rukai) and Bunun (Austronesian, Bunun; cf. Ferrell 1969:68; Radetzky 2004:214).

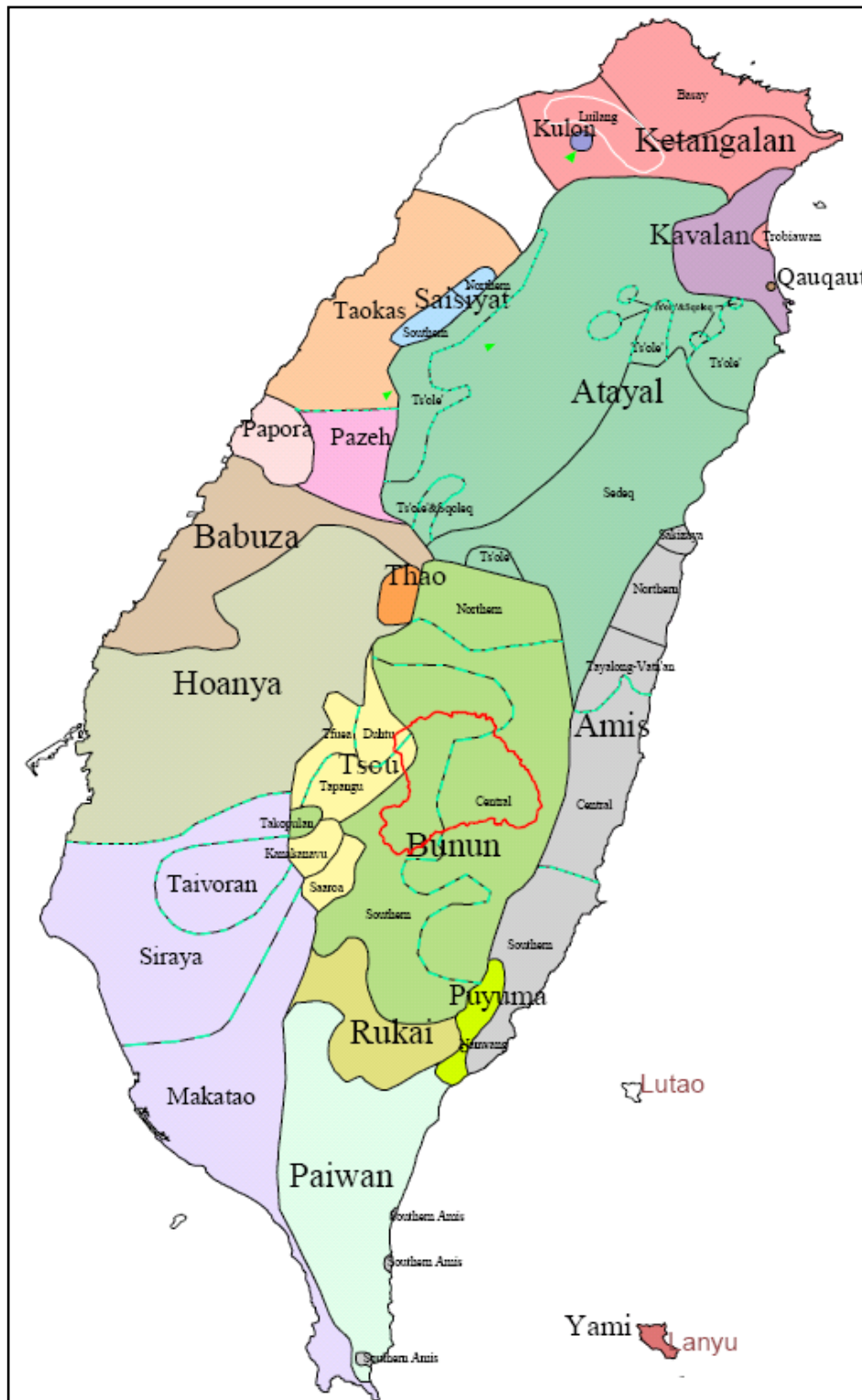
Of the three Tsouic languages two, Kanakanabu and Saaroa, are reported to have interrogative pronominals used indiscriminately as ‘who?’ or ‘what?’. In both languages, the ‘who?, what?’ interrogatives appear to originate in a construction of the structure [‘what?’ + a demonstrative or personal pronominal], literally meaning something like ‘what one?’, i.e. ‘which one? (person or thing)’, which later has been extended to non-selective contexts as both ‘who?’ and ‘what?’. Furthermore, in a few cases the original interrogative part has also been lost, following an evolution similar to that of Italian *cosa?* ‘what?’, that has developed from *che cosa?* ‘what thing?’.

In Section III.4.2.6.1, I will begin by discussing the interrogative pronominals of Kanakanabu. The interrogative pronominals of Saaroa will be examined in Section III.4.2.6.2.

¹¹⁴ I would like to thank Paul Li and Dah-an Ho for sharing with me their data on Kanakanabu and Saaroa. I am also grateful to Marie Yeh, Stacy Teng and Lillian Huang for bibliographical references.

¹¹⁵ If all the Austronesian languages of Taiwan are grouped together as Formosan and opposed to Malayo-Polynesian as Extra-Formosan, Tsouic forms an independent branch of the Formosan part of the Austronesian family. Cf. Li (2006) for a brief overview of the extant classifications of the Austronesian languages of Taiwan.

Map 9. Austronesian languages of Taiwan (adapted from Tsuchida 1983 via <http://gis210.sinica.edu.tw/ysnp/ecai/language.pdf>)¹¹⁶



¹¹⁶ Most of the Austronesian languages of western Taiwan are extinct.

4.2.6.1 Kanakanabu

Kanakanabu is a moribund Tsouic language. The available data on the forms of interrogative pronominals with their respective meanings are summarized in (117).

Kanakanabu¹¹⁷

- | | |
|-------------------------------|--|
| (117) Dah-an Ho (p.c.) | <i>namanaini</i> ‘who?’, <i>naumani</i> ‘what?’ |
| Paul Li (p.c.) | <i>nian</i> or <i>naumani</i> both ‘who?, what?’ |
| Yen <i>et al.</i> (1962-1963) | <i>nein</i> ‘who?, what?’ |
| Ogawa & Asai (1935) | <i>no:mani</i> ‘what?’, no form for ‘who?’ |

Examples (118-122) illustrating the use of two of these forms, *nian* and *naumani*, support their interpretation as ‘who?, what?’ interrogatives.

Kanakanabu

- (118) *naumani kasu?*
 IPW 2SG
 ‘Who are you?’ (Paul Li, p.c., data collected in Min-sheng village)
- (119) *sua iikam miana ya, urupaca-kamu naumani muru-?caŋa?*
 NOM 2PL before TOP use-2PL.AGENTIVE IPW in-marriage
 ‘In your old times, what did you use in the wedding?’ (Dah-an Ho, p.c.)¹¹⁸
- (120) *nian sua iisi?*
 IPW NOM this
 ‘What’s this?’ (Paul Li, p.c., data collected in Min-sheng village)
- (121) *nian sua cau iisa?*
 IPW NOM person that
 ‘Who is that person?’ (Paul Li, p.c., data collected in Min-sheng village)
- (122) *nian manu-musu?*
 IPW child-2SG.POSS
 ‘Who is your child?’ (Paul Li, p.c., data collected in Min-sheng village)

For a language with just a small group of elderly speakers left, the sheer number of forms of interrogative pronominals and discrepancies in the interpretation of their meanings found in the sources are quite remarkable. However, given the availability of examples and an important degree of similarity between the forms from different sources, it seems unreasonable to doubt the reliability of the

¹¹⁷ Yen *et al.* (1962-1963) and Ogawa & Asai (1935) are cited via Ferrell (1969:402, 405).

¹¹⁸ Note that the case marker *sua* is glossed by Paul Li as nominative, while Dah-an Ho describes it as a marker for “both agentive and patient case”. I use the gloss NOM for convenience sake.

sources. When this variation in the sources is considered together with the fact that some forms are relatively long and some look like parts of these longer forms, e.g. *namanaini* vs. *nein*, a conclusion that suggests itself instead is that these interrogatives are phrasal constructions in origin. In what follows, I will first examine the interrogatives *namanaini* and *nein* (Section III.4.2.6.1.1) and then proceed to the form *nian* (Section III.4.2.6.1.2).

4.2.6.1.1 The interrogatives *namanaini* and *nein*

In the case *namanaini*, the final part *-naini* looks very much like a demonstrative or a combination of at least two (and probably three) demonstratives. Consider, for instance, Tsou *eni* ‘this’ (Ferrell 1969:398), Proto Austronesian **i-ni* ‘this’ (Blust *et al.* 2006), Proto Malayo-Polynesian **i-na* ‘that’ (Blust *et al.* 2006), Kavalan *?nay* ‘that’ (East Formosan; Huang & Sung 2005), Proto Northern Luzon **na* ‘medial’ (Northern Philippine; Reid 2006:20, 23), Saaroa ‘oblique’ (Tsuchida 1976:67 cited via Radetzky 2004:215) and the Tfuya Tsou ‘nominative’ noun phrase marker *na* (Huang *et al.* 1999:656, 660).¹¹⁹ If we add to this that interrogative pronominals based on the structure [‘what?’ + a demonstrative or personal pronominal] can be found in various Malayo-Polynesian languages,¹²⁰ we can hypothesize that *namanaini* also goes back to a similar, although probably somewhat more complex structure. Literally, its original meaning was thus something like ‘what one?’, i.e. ‘which one? (person or thing)’, which must have been later extended to non-selective contexts as both ‘who?’ and ‘what?’. In the Kanakanabu variety using *nein* instead of *namanaini*, the interrogative *nein* must have been derived from *nama-naini* through the loss of the original interrogative part, following an evolution similar to that of Italian *cosa?* ‘what?’.

To determine what the original interrogative part in the Kanakanabu interrogatives *namanaini* and *naumani* may have been, let us now compare the two interrogatives to each other and to such interrogative pronominals as Kanakanabu *nuka-naumana* ‘when?’ (Ogawa & Asai 1935 via Ferrell 1969:404), Tsou *cuma* ‘what?’ and *homna* ‘when?’ (Tung 1964 via Ferrell 1969:402, 404), Proto Austronesian **si-ima* ‘who?’ (Blust *et al.* 2006), Puyuma *i-manai* ‘who? (NOM)’ and *a-manai* ‘what? (NOM)’ (Cauquelin 1991:41). Three things may be

¹¹⁹ Consider also the Saaroa affix *-nai* in (134) in Section III.4.2.6.2, left unexplained by Radetzky (2004), which seems to convey some kind of deictic or focus meaning.

¹²⁰ Consider the discussion of the Lampungic interrogative pronominals *api(ya)* and *ɲɔ* in Section III.4.2.3, the Sisingga interrogative *anda* ‘who?, what?’ in Section III.4.2.4, Iwal *ret*, Koronadal Bilaan *det*, Sarangani Bilaan *tan* in Section III.4.2.5, Lower Tanudan Kalinga *dimman* ‘where?’ in Section III.4.2.2.2 (footnote 99).

deduced from this comparison: (i) the original interrogative root in all these interrogatives must be **ma* ‘what?’, (ii) the part preceding **ma* is in all probability a noun phrase marker, either a nominal specifier or case marker, both ultimately going back to a demonstrative root,¹²¹ (iii) the part following **ma* is probably a demonstrative or a frozen combination of demonstrative roots. That is, the original structure of these interrogatives must have been either [noun phrase marker + ‘what?'] or [noun phrase marker + ‘what?’ + a demonstrative or personal pronominal]. In both cases, the distinction between human and non-human meanings of the construction must have been achieved by means of noun phrase markers, which in Austronesian languages are often organized following the opposition “personal” vs. “non-personal” (cf. Section III.4.2.2.1.2 for more details). Yet, it also happens that a noun phrase marker is absent or does not distinguish between personal and non-personal nominals. In this case, [‘what?’ + a demonstrative or personal pronominal] literally meant something like ‘what one?’ (i.e., ‘which one?’) and could either remain ambiguous between human and non-human meanings or become restricted as ‘what?’ (less likely as ‘who?’).

The hypothesis proposed above can be supported by the following arguments. To begin with, note that in the Proto Austronesian reconstruction **si-ima* ‘who?’ both **si* and **i* are most likely the “personal” noun phrase markers, widespread both in Malayo-Polynesian Austronesian languages (cf., e.g., Reid 2006; Reid & Liao 2006:469; François, forthcoming; Section III.4.2.2.1.2) and in Formosan Austronesian languages, e.g. the Mayrinax Atayal and Sguliq Atayal (nominative) “personal” marker *ʔiʔ* (Huang 2006:2) or the Puyuma nominative singular “personal” marker *i* (Cauquelin 1991:42).¹²² Furthermore, the use of a personal noun phrase marker to derive a human interrogative ‘who?’ from the unmarked non-human ‘what?’ is common in Malayo-Polynesian Austronesian languages. For instance, Indonesian has *apa* ‘what?’ vs. *si-apa* ‘who?’ (Malayic; Blust *et al.* 2006) and Binongan Itneg has *ʔano* ‘what?’ vs. *si ʔano* ‘who?’ (Northern Philippine; Reid 1971:161, 163), where both *apa* and *ʔano* are regular reflexes of the two Proto Malayo-Polynesian ‘what?’ interrogatives, **apa* and **-anu* respectively (Blust *et al.* 2006).

Second, in the Puyuma interrogatives *i-manai* ‘who? (NOM)’ and *a-manai* ‘what? (NOM)’ the distinction between human vs. non-human is achieved by means of different case markers. That is, similarly to Kanakanabu *namanaini* and *naumani*, the Puyuma interrogative root *-manai* itself is indifferent to this distinction. It is not implausible then that Puyuma *-manai* may also originate in

¹²¹ Cf. Reid (2002, 2006) on the noun phrase markers in Philippine languages and their demonstrative origins. See also Section III.4.2.2.1.2 for a brief overview.

¹²² Note that of the two markers, **si* and **(ʔ)i*, the second one is likely to be the oldest (cf. Reid 1979:11-16; see also Section III.4.2.2.2).

the structure [‘what?’ + a demonstrative or personal pronominal], with the interrogative part *ma* followed by a demonstrative part *nai*. Compare, in this respect, the Kavalan demonstrative *?nay* ‘that’ mentioned above (East Formosan; Huang & Sung 2005). That is, the Puyuma interrogatives *i-manai* and *a-manai* may go back to the structure [case marker + ‘what?’ + a demonstrative or personal pronominal] literally meaning ‘what one?’ (i.e., ‘which one?’), person or thing depending on the case marker.

When we now turn back to the Kanakanabu forms *namanaini* and *naumani*, we should first of all recall that in the two other Tsouic languages, Saaroa and (Tfuya) Tsou, *na* occurs as an “oblique” and as a “nominative” noun phrase marker respectively (Tsuhida 1976:67 cited via Radetzky 2004:215; Huang *et al.* 1999:656, 660). Furthermore, another variety of Tsou, which has a different nominative marker *’o*, also appears to sometimes use a combination *na ’o* to mark nominative NPs, as in (123a) vs. (123b).¹²³

Tsou (Li 2006:5)

- (123) a. *m-oso m-imo to emi ’o ic’o*
 AFOC-AUX AFOC-drink OBL wine NOM that
 ‘That person has drunk wine.’
- b. *oh ta im-a (na) ’o emi*
 PFOC.AUX [OBL]3SG drink-PFOC NOM wine
 ‘The wine has been drunk by him’

Example (124) illustrates the use of the nominative marker *na* in Tfuya Tsou.

Tfuya Tsou (Huang *et al.* 1999:660)

- (124) *(zou) cuma na eni?*
 be IPW NOM this
 ‘What is this?’

What is important here is that neither *na* nor *(na) ’o* distinguish between personal vs. non-personal (or human vs. non-human, animate vs. inanimate), which may explain why *namanaini* and *naumani* are indifferent to this distinction as well.¹²⁴

¹²³ Li (2006:5) himself leaves *na* in (123b) unglossed.

¹²⁴ Example (124) above is also of particular interest because the copula *zou* ‘be’ is likely to be of a demonstrative origin (compare, e.g., the Kavalan demonstrative *zau* ‘this’, Huang & Sung 2005), i.e. of the same origin as the noun phrase markers *na* or *(na) ’o*. This suggests that the original construction of the interrogatives *namanaini* and *naumani* might have been structurally similar to (124).

4.2.6.1.2 The interrogative *nian*

The Kanakanabu interrogative *nian* ‘who?, what?’ may have various origins. The hypothesis I found most promising is that *nian* originates in something like **naumani-an* (or **namanaini-an*) in the same way as *nein* has presumably developed from *nama-naini* (cf. Section III.4.2.6.1.1). The suffix *-*an* in the hypothetical form **naumani-an* is in all probability the frozen accusative/locative case suffix -*an* common in Formosan languages. The same suffix -*an* is also frequently found on verbs both in Formosan and Extra-Formosan languages as the so-called “locative focus” marker or “patient focus” (sometimes, “non-agent focus”) marker.¹²⁵ Examples (125-127) illustrate the use of the verbal and nominal suffix -*an* in several Formosan languages.

Northern Paiwan (Huang *et al.* 1999:658)

- (125) a. *tima su-k <in> isədam-an tua paysu?*
 NOM.IPW 2SG.GEN-borrow<PFV>-LFOC ACC money
 ‘Who did you borrow money from?’
 b. *nima alak a su-k <in> atjəŋəray-an?*
 GEN.IPW child NOM 2SG.GEN-like<PFV>-LFOC
 ‘Whose child do you like?’

Central Amis (Huang *et al.* 1999:659)

- (126) a. *cima ku ma-ulah-ay ci panay-an?*
 NOM.IPW NOM AFOC-like-AFF ACC PROP-ACC
 ‘Who likes Panay?’
 b. *nima wawa ku ka-ulah-an isu?*
 GEN.IPW child NOM PFOC-like-PFOC 2SG.GEN
 ‘Whose child do you like?’
 c. *cimanan kisu pa-fli? tuni cuʔaʔ?*
 IPW.ACC 2SG.NOM CAUS-give this book
 ‘To whom will you give this book?’

Kavalan

- (127) a. *pa-q-azin = iku qanyau-an*
 CAUS-AFF-tell=1SG.NOM 3PL-ACC
 ‘I recognized them’ (Huang & Sung 2005:2)

¹²⁵ In descriptions of Austronesian languages, the term “focus”, such as “agent focus”, “patient focus”, “locative focus”, etc., refers to verbal affixes that have something to do with the valence of the verb.

- b. *pa-q-azin = iku* *ta-qanyau-an*
 CAUS-AFF-tell=1SG.NOM LOC-3PL-LOC
 ‘I recognized them’ (Huang & Sung 2005:2)
- c. *zuit-an* *ya* *taquq* *?nay ta-paRing-an*
 hang-NAFOC NOM chicken that LOC-tree-LOC
 ‘He hung a chicken on the tree’ (Huang & Sung 2005:4)

Note the Central Amis accusative interrogative *cimanan* ‘who?’ in (126c). However, an even better parallel to the Kanakanabu interrogatives *naumani* and **naumani-an* can be found in Mayrinax Atayal. This language has two interrogative pronominals, *nanu?* ‘what?’ and *nanuwan* ‘what?’, which are in free variation, (128a) vs. (128b), except that only *nanuwan* can “serve as a predicate appearing in sentence initial position” (Huang 1996:275), as in (128c).

Mayrinax Atayal

- (128) a. *si-βainay cu? nanuwan nku? kanairil ku? ?ulaqi??*
 BFOC-buy ACC IPW GEN.REF woman NOM.REF child
 ‘What did the woman buy for the child?’ (Huang 1996:274)
- b. *si-βainay cu? nanu? nku? kanairil ku? ?ulaqi??*
 BFOC-buy ACC IPW GEN.REF woman NOM.REF child
 ‘What did the woman buy for the child?’ (Huang 1996:274)
- c. *nanuwan ku? β<in>ainay ni? yaya? ?i? isu??*
 IPW NOM.REF buy<PFV.PFOC> GEN mother ACC 2SG.NEU
 ‘What did Mother buy for you? (lit.: ‘What (was) it that Mother bought for you?’)’ (Huang 1996:272)

In all probability, the final *-an* in *nanuwan* represents the fossilized case suffix *-an*, already familiar to us. Interestingly, similarly to Kanakanabu, Mayrinax Atayal does not use *-an* elsewhere as a case suffix, whereas other Atayal varieties, such as Wulai Atayal and Plngawan Atayal, still recur to it in the case paradigms of their personal pronouns (cf. Huang 2006).

4.2.6.2 Saaroa

Saaroa is a moribund Tsouic language spoken by some 15 elderly people in the T’aoyüan and Kaochung villages of Kaohsiung County (Radetzky 2004:213). The available data on the forms of interrogative pronominals with their respective meanings are summarized in (129).

Saaroa¹²⁶

- (129) Tsuchida (1969), Paul Li (p.c.) *ŋaʔaɪsa* ‘who?, what?’
 Radetzky (2004) *aʔaɪsa* ‘who?’, no form for ‘what?’
 Yen *et al.* (1962-1963) *ŋaʔaɪsi* ‘who?’, no form for ‘what?’

Examples (130-134) illustrating the use of the interrogative (*ŋ*)*aʔaɪsa* support its interpretation as a ‘who?, what?’ interrogative.

Saaroa

- (130) *ŋaʔaɪsa pi-salia?*
 IPW own-house
 ‘Who owns a house?’ (Paul Li, p.c.)
- (131) *ŋaʔaɪsa ʔiuvura na iʔau?*
 IPW gave OBL 2SG
 ‘Who gave it to you?’ (Paul Li, p.c.)
- (132) *ŋaʔaɪsa ʔiuvura-isa cucuʔu iʔau?*
 IPW gave-3.GEN person 2SG
 ‘What did the person give you?’ (Paul Li, p.c.)
- (133) *ŋaʔaɪsa kanaʔa?*
 IPW this
 ‘What’s this?’ (Paul Li, p.c.)
- (134) *are, aʔaɪsa ka umu tabiʔibiʔi-ku-nai?*
 oh IPW NS eat banana-1SG-AFF
 ‘Oh, who has eaten my banana?’ (Radetzky 2004:229)

I believe that the Saaroa interrogative (*ŋ*)*aʔaɪsa* ‘who?, what?’ is due to an evolution similar to the one that has resulted in Kanakanabu in the ‘who?, what?’ interrogatives *nein* and *nian* (Section III.4.2.6.1). That is, in all probability, (*ŋ*)*aʔaɪsa* goes back to the structure [‘what?’ + a demonstrative or personal pronominal], literally meaning ‘what one?’, i.e. ‘which one? (person or thing)’, with a subsequent loss of the initial interrogative part and extension to non-selective contexts. The following evidence can be adduced in support of this hypothesis.

First, the *-aɪsa* part of (*ŋ*)*aʔaɪsa* strongly resembles the third person (singular or plural) anaphoric pronominal ‘he, she, it, they’ and distal deictic ‘that’ *iʔaɪsa* (Tsuchida 1969 via Ferrell 1969:399). Note also that all free anaphoric pronominals in Saaroa begin with *iʔa-*, while *-isa* of *iʔaɪsa* must be cognate to Kanakanabu *iisa* ‘that’, as in (121) above. Kanakanabu also has a proximate demonstrative *iisi* ‘this’, as in (120) above, which may explain the form *ŋaʔaɪsi* in

¹²⁶ Tsuchida (1969) and Yen *et al.* (1962-1963) are cited via Ferrell (1969:402, 405).

Yen *et al.*'s (1962-1963) data, (129).

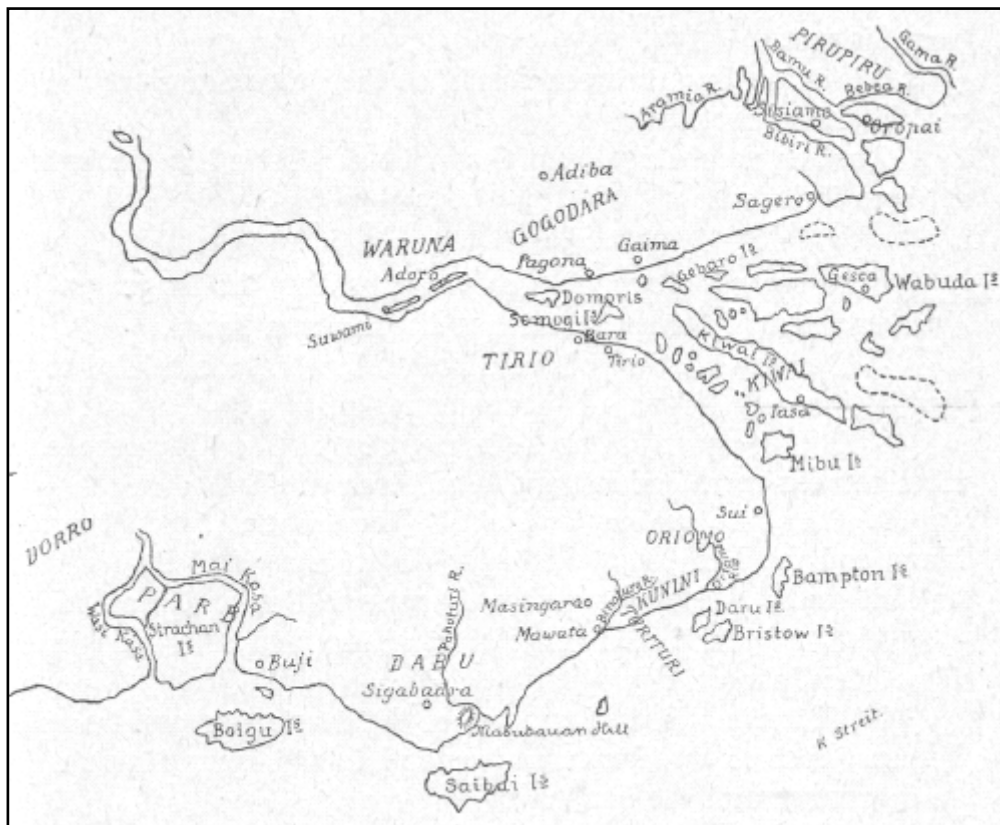
Second, the lost initial interrogative part of *(ŋ)ałaisa* has been preserved in the Saaroa interrogatives *ki:łaumajə* 'when? (in the past)' and *cułaumajə* 'when? (in the future)' (Tsuchida 1969 via Ferrell 1969:404). In this respect, compare also Kanakanabu *nukanaumana* 'when?' (Ogawa & Asai 1935 via Ferrell 1969:404) and Tsou *homna* 'when?' (Tung 1964 via Ferrell 1969:404). Apparently, the earlier interrogative part of the Saaroa interrogative *(ŋ)ałaisa* has undergone the following development **ma-na-* > **ma-ŋa-* > *(ŋ)a-*. The original interrogative root here happens to be the same interrogative root **ma* 'what?' already discussed in Section III.4.2.6.1.1.

Finally, besides being plausible formally and semantically the hypothesis proposed squares well the similar developments reconstructed for the Kanakanabu 'who?, what?' interrogatives *nein* and *nian* (Section III.4.2.6.1).

5 New Guinea: Kiwaian languages

Kiwaian languages are primarily spoken on the coast of Western Province of Papua New Guinea, in particular in the deltas of the Fly and Bamu Rivers, cf. Map 1.

Map 1. Some languages spoken in the coastal regions of Western Province, Papua New Guinea (Riley & Ray 1930:173)



According to the Ethnologue, the Kiwaian family belongs to the Trans-Fly-Bulaka River subphylum of the Trans-New Guinea phylum.

There appears to be at least one Kiwaian idiom, the Pirupiru dialect of the Bamu language, with an interrogative used indiscriminately as ‘who?’ or ‘what?’. Table 1 summarizes the interrogatives ‘who?’, ‘what?’ and ‘which one?’ in some Kiwaian languages. As can be observed from a comparison between the Pirupiru Bamu and the Sisiamé Bamu interrogatives ‘who?’ and ‘what?’, the Pirupiru interrogative *etura* ‘who?, what?’ most likely results from a conflation of the earlier interrogatives **etura* ‘who?’ and **etoura* ‘what?’ through a (more or less accidental) contraction **ou > u* in the non-human interrogative **etoura > etura*. Note in this respect that the ‘what?’ interrogatives in the other Kiwaian idioms in Table 1 have *a* where Sisiamé Bamu has *ou*, which apparently has

Table 1. Interrogatives ‘who?’, ‘what?’ and ‘which one?’ in some Kiwaian languages (based on Riley & Ray 1930:193; Ray 1933:21-23)

	‘who?’	‘what?’	‘which one?’
Bamu, Pirupiru dialect	<i>etura</i>		<i>bidanara</i>
Bamu, Sisiamé dialect	<i>etura</i>	<i>etoura</i>	<i>bedanara</i>
Bamu, Oropai dialect	<i>etura</i>	<i>bedara</i>	<i>bedanara</i>
Southern Kiwai, Iasa dialect	<i>betu</i> (SG), <i>betugoto</i> (DU), <i>bedigo</i> (PL)	<i>beda, (e)beta</i>	<i>bedana</i>
Southern Kiwai, Turituri dialect	<i>boturo</i>	<i>beda</i>	<i>bedana</i>
Southern Kiwai, Doumori dialect	<i>beiro</i>	<i>bedaro</i>	<i>budonuunaro</i>
Waboda	<i>o(ra)turo</i>	<i>dabaro,</i> <i>dabanamora</i>	<i>dabanamora</i>

helped to preserve the distinction between the interrogatives ‘what?’ and ‘who?’ there. Therefore, it might be interesting to briefly discuss in the rest of the present section the development of this *ou* in the Sisiamé Bamu (and earlier Pirupiru Bamu) interrogative *etoura*.

To begin with, note that the interrogatives in Table 1 are clearly polymorphemic in origin. The element that can be separated most easily is the final *-ro/-ra*. Thus, in Iasa Southern Kiwai, *-ro* is a “particle [...] often used to distinguish the most important pronoun in the sentence [including] interrogatives” (Ray 1933:4, 21-23).¹ Other such particles that often occur on the interrogatives in Iasa Southern Kiwai are *-go* and *-ia*, for instance (Ray 1933: 21). The syllables *-tu* and *-di* in the ‘who?’ interrogatives are in all probability related to the Iasa Southern Kiwai words *dubu* ‘man’, *didiri* ‘men’, *(a)rubi* ‘people’ (cf. Ray 1933:4, 10-11). The syllable *-ta/-da* in the ‘what?’ interrogatives is cognate to the Iasa Southern Kiwai pronominal *ata* ‘another’ (cf. Ray 1933:8). The elements *-na*, *-nuuna* and *-(ba)namo* in the ‘which one?’ interrogatives must be related to the Iasa Southern Kiwai word *nuunumabu* ‘thing’ (1a), which can also be “abbreviated to” *nuuna* (1b) or *-na* (1c) (Ray 1933:11).

Iasa Southern Kiwai (Ray 1933:11)

- (1) a. *kopirawa nuunumabu*
hidden thing
‘a secret’

¹ In Table 1, the final *-ro/-ra* in the interrogatives of the idioms other than Iasa Southern Kiwai may be optional as well. However, Riley & Ray’s (1930) wordlist does not go into such details.

- b. *obo kiaputi nuuna*
 water to.draw thing
 ‘something to draw water with’
- b. *wade-na*
 good-thing
 ‘a good thing’

The initial *(e)be-* or *be-/bi-* in the interrogatives in Table 1 must be the original attributive interrogative root ‘which [N]?, what [N]?’. Thus, literally, ‘who?’ in Kiwai and Bamu used to mean something like ‘what man?’ and ‘what?’ something like ‘what another?’ (i.e., ‘what one?’). The *-ou-* in the Sisiame Bamu (and the earlier Pirupiru Bamu) interrogative *etoura* ‘what?’ can be explained easily if we presume that like the Waboda interrogative *dabanamora* ‘which one?, what?’, *etoura* used to be a ‘which one?’ interrogative, **e(be)tanuunara* (or the like; compare also the Doumori Southern Kiwai *budonuunaro* ‘which one?’).

6 Australia¹

6.1 Introduction²

Australian languages are usually subdivided into two blocks: Pama-Nyungan languages, spread almost all over the continent, and the so-called non-Pama-Nyungan languages, concentrated in Northern Australia (excluding the Cape York Peninsula).³ The received view is that Pama-Nyungan languages are much less diverse than non-Pama-Nyungan languages. It is also commonplace to

¹ I am grateful to Mark Harvey for his comments on a previous version of this section.

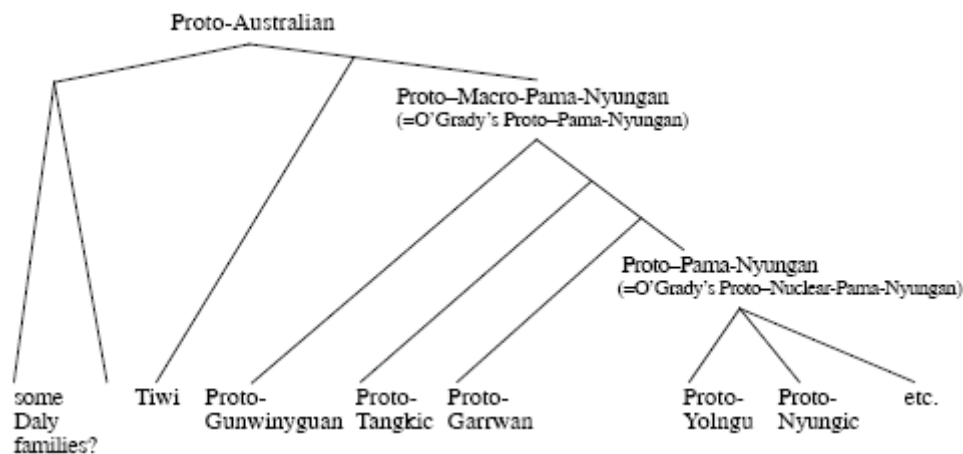
² Whenever possible, I have uniformized the spelling of examples and single forms in this section. Here is the list of the most important orthographic conventions adopted for the representation of Australian data:

- (i) The majority of the Australian languages, especially Pama-Nyungan, do not distinguish between voiced and voiceless stops. However, different spellings may prefer to spell one and the same stop phoneme as either “voiced” or “voiceless”. Some orthographies use both options, but in different contexts. Many non-Pama-Nyungan languages have two series of stops, usually described in terms of *fortis* vs. *lenis* or *geminate* vs. *simple* (cf. e.g. Butcher & Reid 1989; Baker 1999). Here again, orthographies differ in the way they represent the two series. Usually, I follow the original source in its use of the “voiced” and “voiceless” stops in the spelling.
- (ii) I use *j* instead of *dj* or *dy* for the “voiced” lamino-palatal stop, other lamino-palatal stops are rendered by digraphs as *ny*, *ly* and *ty*. The palatal glide is rendered with *y*.
- (iii) Following the common practice, *h* after a consonant marks it as lamino-(inter)dental, e.g. *th*, *dh*, *nh*, and *lh*. Note that in some (non-Pama-Nyungan) orthographies *h* marks the glottal stop ʔ. In the latter case, whenever known, I use the glottal stop sign instead of *h*.
- (iv) Following the common practice, a single *r* normally stands for the retroflex rhotic. Before another consonant, *r* marks the latter as retroflex, e.g. *rd*, *rt*, *rn*, *rl*. Note, however, that double *rr* is not a retroflex but an apico-alveolar rhotic. Retroflexion is also sometimes marked with a dot under the consonant, e.g. *ṛ* for *rd*.
- (v) When two palatal, laminal or retroflexed consonants occur in a sequence, some orthographies mark the palatal, laminal or retroflexed realizations only once, e.g. *nj* may be used for *ny* + *j*, *nth* may be used for *nh* + *th*, *rnd* may be used for *rn* + *rd*.
- (vi) By default, the digraph *ng* has the same value as *ŋ*, although sometimes it may represent a combination of *n* or *ŋ* and *g*. To avoid complications or for the sake of uniformity, I often use *ng* even when it is known to represent the velar nasal.

³ Note that the long extinct languages of Tasmania are usually left outside of the scope of this classification due to the lack of data.

consider Pama-Nyungan languages as a single genetic unit,⁴ whereas “existing classifications of non-Pama-Nyungan languages postulate some twenty-seven coordinate language families, with no-higher order subgroupings, and many of the families having a single member” (Evans 2003a:11). Cf. Map 1 for the distribution of the non-Pama-Nyungan language families.⁵ According to some recent attempts to apply the comparative method to non-Pama-Nyungan languages, there is enough evidence pointing to “the relatedness of nearly all non-Pama-Nyungan languages” (Evans 2003a:17). Furthermore, it is suggested that Pama-Nyungan and most non-Pama-Nyungan languages are also related. In this perspective, Pama-Nyungan languages are very likely to represent “a relatively recent daughter node within a larger Stammbaum containing most extant Australian languages” (Evans 2003a:9). A version of such a possible genetic tree uniting most Australian languages is illustrated in Figure 1. For a discussion and further references see, for instance, Evans (2003a).

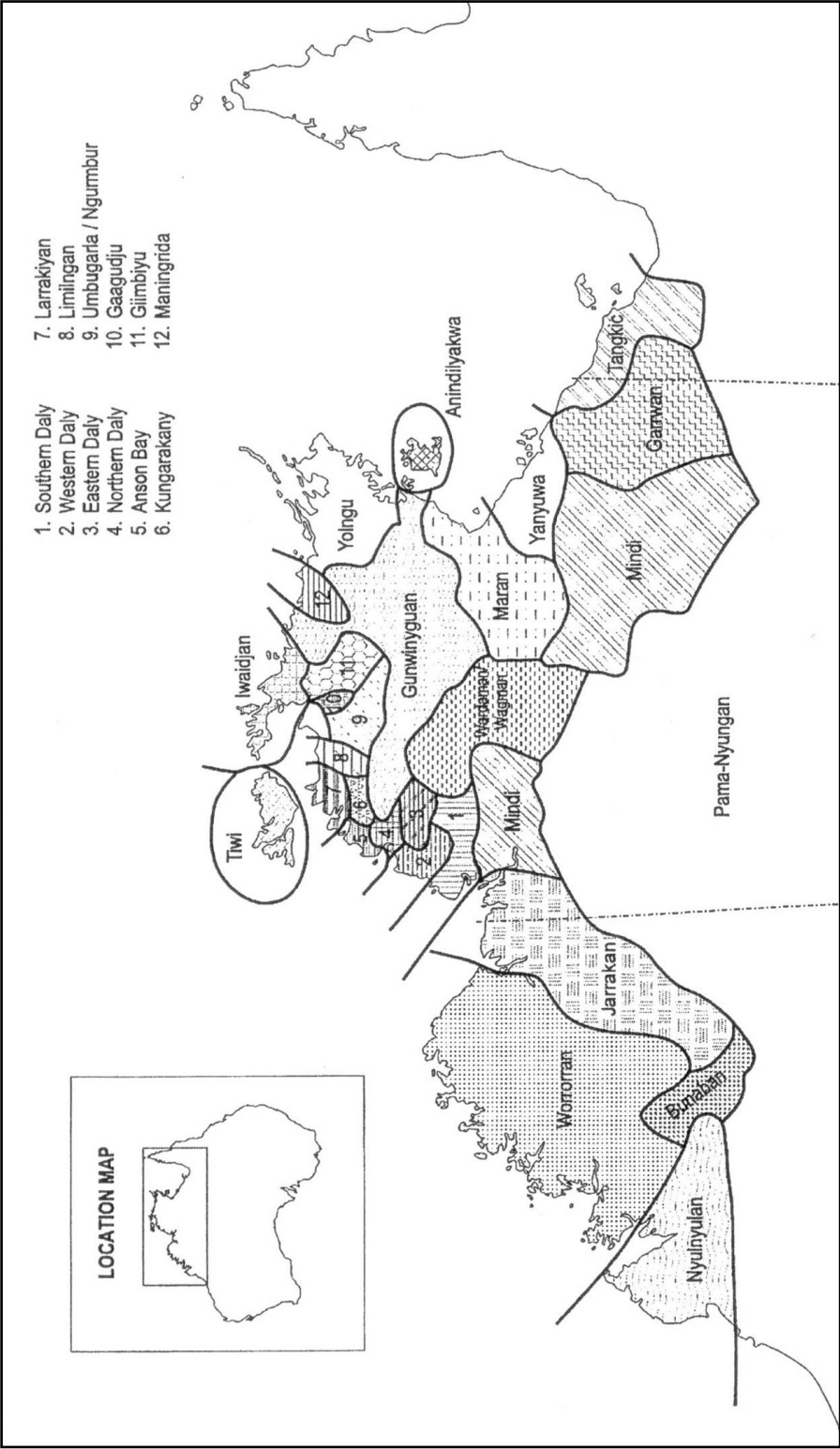
Figure 1. “The Pama-Nyungan offshoot model, with Pama-Nyungan as an offshoot sharing immediate ancestry with some non-Pama-Nyungan groups, after O’Grady (1979), Evans & Jones (1997)” (Evans 2003a:10)



⁴ Although the Pama-Nyungan hypothesis seems to enjoy support of the majority of the Australianists, it has been strongly opposed by R.M.W. Dixon, mostly within the framework of his “Punctuated Equilibrium” model (1980, 1997, 2002). For a criticism of Dixon’s statements and further references to the more mainstream literature see, for instance, Evans (2003a).

⁵ Yolngu and Yanyuwa are Pama-Nyungan enclaves. In what follows, I will use the names of the non-Pama-Nyungan language families as presented on this map. The name of the Mindi language group is spelt further as Mirndi.

Map 1. Non-Pama-Nyungan language families (Evans 2003c:2)



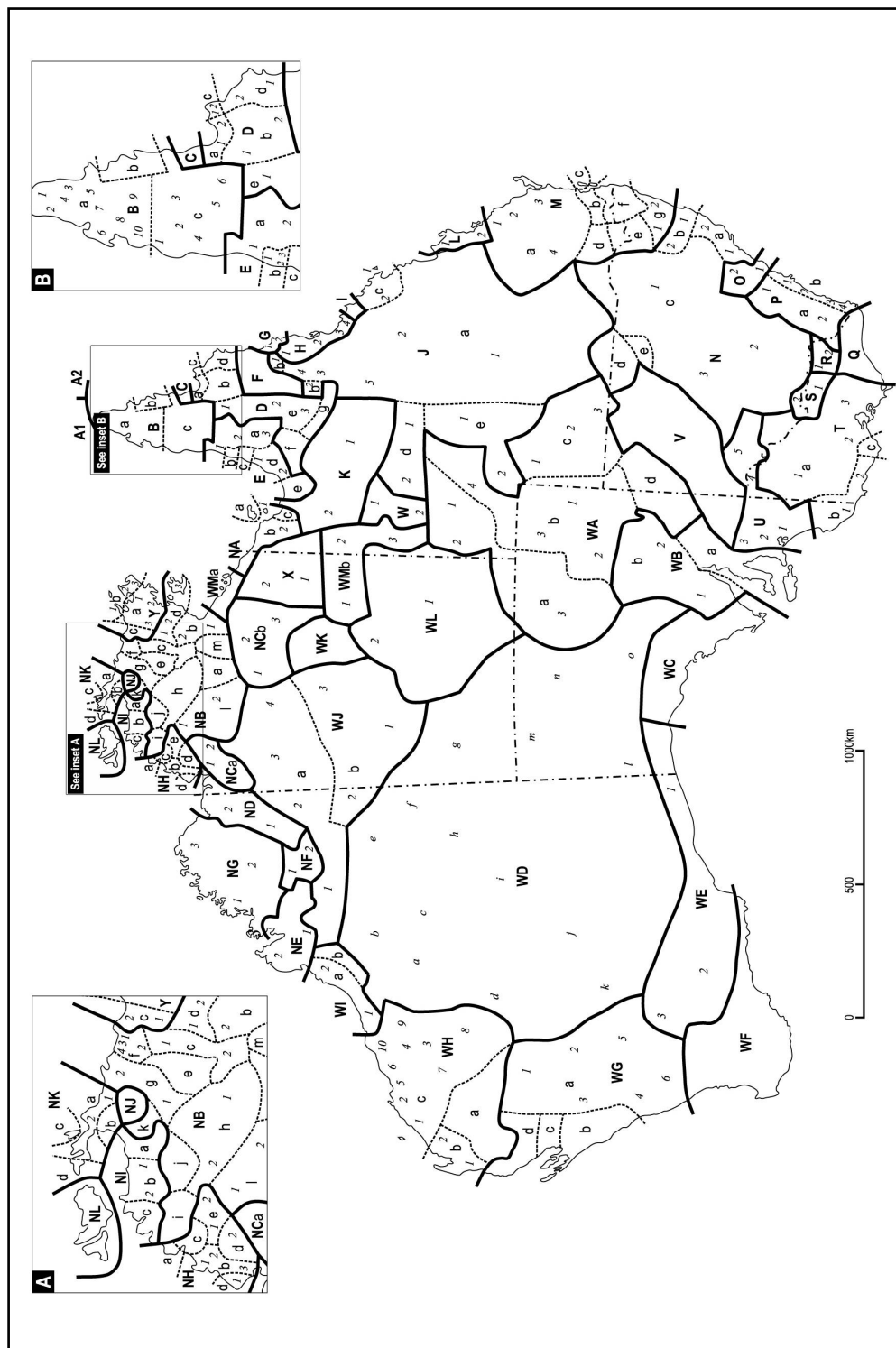
The classification differing most radically from the one presented in Figure 1 has been advanced by Dixon (2002), who regards “Australian languages as making up a large linguistic area” (2002:xxiv), where long-time contact and extensive diffusion of linguistic traits must have erased all traces of any earlier higher-level genetic subgrouping. Dixon prefers to organize all Australian languages into fifty smaller groups, “for ease of reference” labelled A-Y, WA-WM and NA-NL, with lower case letters used to mark “groups within groups” and numbers for languages (2002:xxv), as illustrated on Map 2. Some of the groups are “tentatively identified as low-level genetic subgroups”, some “as small linguistic areas”, while the remaining groups “simply consist of languages grouped together on a geographical basis” (2002:xxv). For a critical discussion of Dixon’s (2002) classification see Evans (2005). Here, I prefer to follow the more traditional classification, but for purpose of reference I will also often use Dixon’s (2002) group and language labels.

In what follows, I will first provide a general overview of the geographic and genetic distribution of the Australian languages that (may) allow for the lack of differentiation between ‘who?’ and ‘what?’, Section III.6.2. In Section III.6.3, I will briefly present some common functional and formal patterns of the interrogative pronominals in Australian languages that may be relevant for accounting for the lack of differentiation between ‘who?’ and ‘what?’ in certain languages. Then, in Section III.6.4 the relevant languages and forms will be considered in more detail. Finally, in Section III.6.5 some concluding remarks will be made.

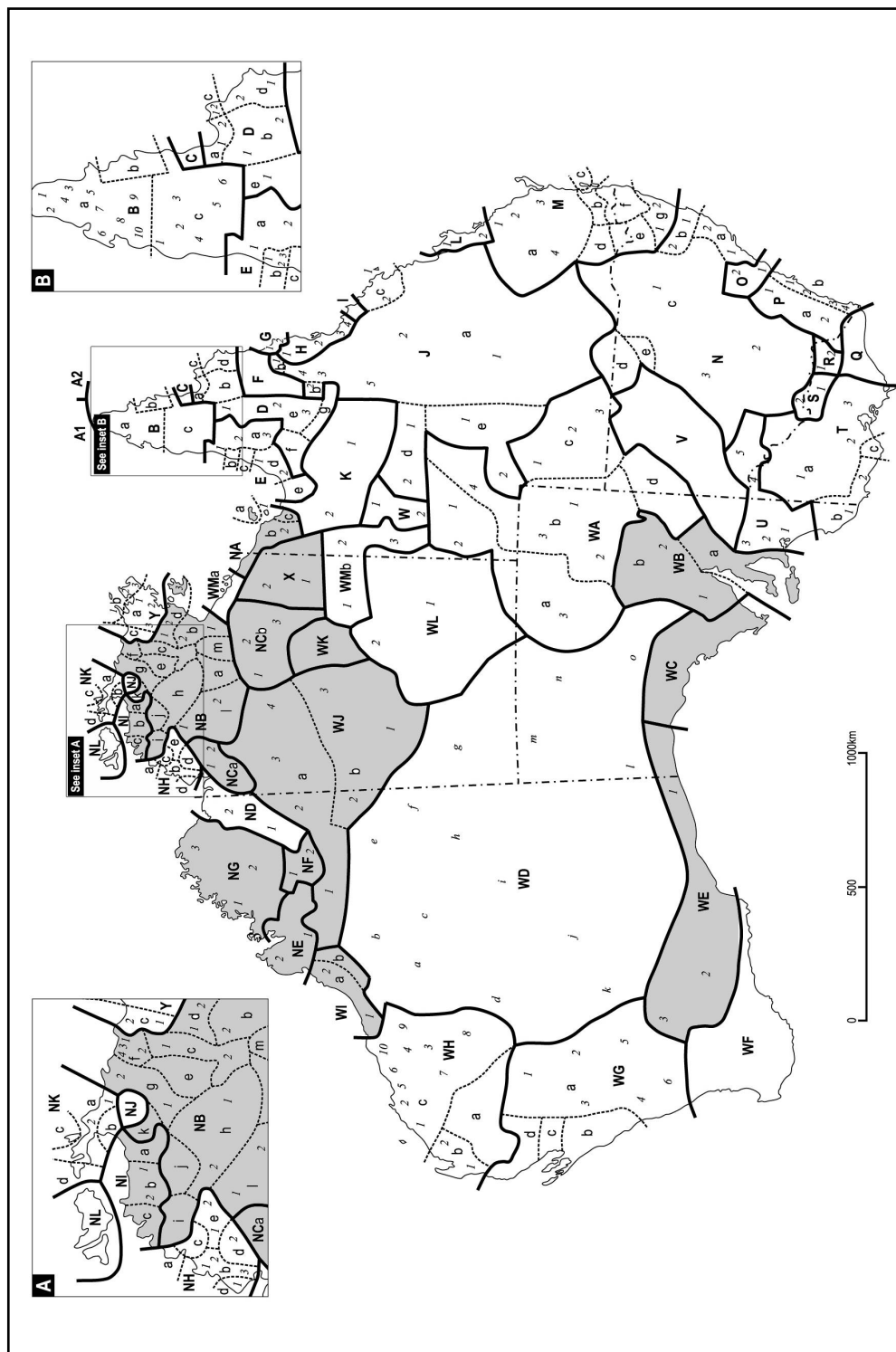
6.2 Australian languages with ‘who?, what?’ interrogatives: distribution

Languages with interrogative pronominals used indiscriminately as ‘who?’ or ‘what?’ can be found both among the Pama-Nyungan and the non-Pama-Nyungan languages. According to Dixon (2002:328), “in about twenty languages a single form is used for both ‘who’ and ‘what’ – this applies to languages from groups X, WB, WC, WE, WI-WK, NA-NC, NE-NG and NI ([...] there are no examples from the eastern third of the continent)”, cf. Map 3. However, a few remarks must be made about this statement. First, depending on what counts as a language and what as a dialect, Dixon’s “twenty” may easily be thirty or even more for others. Dixon uses lack of mutual intelligibility as the main criterion in distinguishing languages from dialects, while other linguists may use the term language “in the political sense”, when “tribal dialects” are called languages, since “for the [Aborigines] themselves it is the tribal dialect (= political language) that has a name” (Dixon 2002:xxiv, 4-7). For instance, Dixon

Map 2. Australian language groups and languages following Dixon (2002:xxviii)



Map 3. Australian language groups with languages where “a single form is used for both ‘who’ and ‘what’” (based on Dixon 2002:328)



distinguishes only two languages in his “Fitzroy River subgroup” NE (traditionally called Nyulnyulan languages), NE1, with Njigina, Warrwa, Yawuru and Jukun dialects, and NE2, with Baardi, Njul-Njul, Djabirr-Djabirr, Ngumbarl and Nimanburru dialects, while Stokes & McGregor (2003:29-31) prefer to distinguish ten: Bardi (with Mainland and Island Bardi dialects), Jabirrabirr, Jawi, Jukun (this maybe a dialect of Yawuru), Ngumbarl, Nimanburru, Nyikina (with Big and Small Nyikina dialects), Nyulnyul (with Coastal and Inland Nyulnyul), Warrwa, Yawuru (with Julbaya and Marangan dialects).⁶

Second, it should be kept in mind that not all languages from the groups “X, WB, WC, WE, WI-WK, NA-NC, NE-NG and NI” use a single form for both ‘who?’ and ‘what?’, since the total number of languages in these groups, 63, by far exceeds 20.

Third, it seems that when Dixon writes “a single form is used for both ‘who’ and ‘what?’” he has at least three different things in mind: (i) a single word used for both ‘who?’ and ‘what?’, as in Walmajarri *ngana* (Pama-Nyungan; WJa1; Dixon 2002:328), (ii) the words for ‘who?’ and ‘what?’ are synchronically derived from one root, such as Ngandi (Gunwinyguan; NBd1)⁷ *-nyja* (Capell 1942:385-386; Heath 1978), (iii) the words for ‘who?’ and ‘what?’ are (in Dixon’s view, it seems) diachronically related through the use of one and the same root, as in the (non-Pama-Nyungan) Bunaban languages Bunuba and Gooniyandi (group NF), for instance.⁸ However, if the existence of a synchronic

⁶ As can be readily observed, the spellings of the language names used by Dixon (2002) differ in several cases from those used by Stokes & McGregor (2003). For various reasons, such variation in the spellings of the names of Australian languages is very common. As I am not in position to uniformize them, I just selected one variant and tried to use it consistently throughout for the languages which are mentioned more than once. In most other cases, I have preferred to preserve the orthography of the sources. However, I may sometimes modify the spelling or add another common variant, for instance when the language is discussed in more detail or when the two common spellings are too different to be immediately recognizable as referring to one and the same idiom.

⁷ Dixon (2002:329) explicitly mentions Ngandi as an example of the language “where ‘who’ and ‘what’ fall together”, although he does not provide the forms.

⁸ Thus, Bunuba has *ngunda* ‘who?’ (Dixon 2002:332) and *ngaanyi* ‘what?’ (Rumsey 2000:74). At least the first form *ngunda* ‘who?’ is mentioned by Dixon (2002:332) among the possible reflexes of the interrogative pronominal **nga(:)n-*, even though he remarks that “it is not obvious at all that all of these forms [i.e., all the forms he adduces as the possible reflexes of **nga(:)n-*] are genetically related”. The Gooniyandi interrogative pronominals, *ngoorn-doo* ‘who?’ (or *ngurndu* in a different spelling) and *jaji* ‘what?’ (McGregor 1990:147-148), do not resemble each other at all. The only other Gooniyandi interrogative that may be considered as

or diachronic relation between the roots for ‘who?’ and ‘what?’ has indeed been the reason for including some of the language groups in Dixon’s list, it seems very strange that many more similar languages from other groups have not been included as well. For instance, in the Girramay dialect of Dyirbal discussed by Dixon (2002:330-331) ‘who?’ and ‘what?’ are also historically based on one root *wany-*,⁹ and the same is claimed by Dixon (1977:195) to be true for Yidiny (G2).¹⁰

Finally, I found at least two Pama-Nyungan languages which appear to have an interrogative ‘who?, what?’, but belong to a group not included in Dixon’s list. The languages at issue are Yulparija/Yulparitja (WDb) and Pintupi (Wdf), both belonging to the so-called Western Desert (or Wati) group (WDb in Dixon’s classification). In addition to Yulparija and Pintupi, two more Pama-Nyungan “languages” are mentioned by Schmidt (1919) that may prove to allow (or to have allowed) for a lack of differentiation between ‘who?’ and ‘what?’ and that do not belong to the groups included in Dixon’s list.¹¹ The first language is Lurija/Luritja (WDg), which like Yulparija is a Western Desert idiom. The second one is a group of languages or a dialect cluster that Schmidt calls “North Narrinyeri”, already extinct long time by now. As far as I can judge from the map of Australian languages in Schmidt (1919), his “North Narrinyeri” seems to correspond to Dixon’s languages U2 (Ngayawang), U3 (Yuyu or Ngarrket), U4 (Keramin, or Kureinji, etc.), and U5 (Yitha-Yitha, Dardi-Dardi).

All in all, I have managed to check around three quarters of the sixty three languages belonging to Dixon’s groups X, WB, WC, WE, WI-WK, NA-NC, NE-NG and NI.¹² In addition, I tried to consult data on at least one language from most other groups, although less consistently in the eastern part of the

being similar at least to *ngoorndoo* is *ngoonyoo* (= *ngunyu*) ‘where?, which one? (person or thing)’.

⁹ According to Dixon (2002:330-331), the Girramay interrogative *wanya* ‘what?’ is originally the “S form” of the interrogative ‘who?’, which in present-day Girramay has the ergative form *wanyju* and the absolutive form *wanyunya*.

¹⁰ The forms of the Yidiny interrogative pronominals *wanyju* ‘who?’ and *wanyi* ‘what?’, almost identical to those of Girramay, are also mentioned by Dixon (2002:330).

¹¹ Schmidt (1919) also gives a single ‘who?, what?’ interrogative *ɲanna* for what he calls “middle languages” (“Mittelsprachen”) of his “south-western group” of Australian languages. However, he remarks that the forms are uncertain. Furthermore, it is not clear what his “middle languages” are exactly. Thus, they may correspond to Dixon’s group WE, which is included in Dixon’s list of languages with a single form used as both ‘who?’ and ‘what?’, but just as well they may cover WE, WH, WG, WI and the majority of the languages of the group WD altogether.

¹² A given “language” has been counted as checked, when it was possible to find information for at least one of its “dialects”. Sometimes, more than one “dialect” has been checked.

continent.¹³ I have summarized on Map 4 the languages that I found to allow for a lack of differentiation between ‘who?’ and ‘what?’ (marked with a grey fill), as well as a few languages for which the available data were inconclusive (marked with a light grey diagonal pattern). For the sake of comparison, similarly to the Map 3, Map 4 is based on Map 2, Dixon’s (2002:xxviii) “master map of language groups and languages” of Australia. Note that where I added group-internal language borders, these are rather schematic. Dixon’s abbreviations used in Map 4 are explained in (1) for the relevant languages. Thus, languages that have proven to allow for a lack of differentiation between ‘who?’ and ‘what?’ are summarized in (1a), while the rest, for which the data have been inconclusive, in (1b). If a “language” has several “dialects” and only one dialect has been checked, this dialect is highlighted in italics.

In total, there are, in Dixon’s terms, 16 “languages” that have proven to allow for a lack of differentiation between ‘who?’ and ‘what?’. Depending on how one defines a language, this number may increase to almost 30 or even more. As can be observed on Map 4, most Australian languages with ‘who?, what?’ interrogatives can be organized in three areas, supplemented by a few sporadic instances elsewhere in the non-Pama-Nyungan part of the continent. Almost all “languages” for which the available data were inconclusive (12 or somewhat more, depending on what counts as “language”) are also immediately adjacent to one of the three areas.

Remarkably, the two northern areas are almost contiguous, with just one or two languages separating them. These two areas are also particularly interesting because each includes both Pama-Nyungan and non-Pama-Nyungan languages. Except Yulparija, the Pama-Nyungan languages at issue all belong to the so-called Ngumpin-Yapa subgroup (Dixon’s WJ group), which are believed to have spread from the south(-east) to their present location only relatively recently, probably some two to three thousand years ago (McConvell 2005). This and the fact that ‘who?’ and ‘what?’ are not differentiated only in the most western and the most eastern Ngumpin-Yapa languages, bordering two clusters of non-Pama-Nyungan languages with the same feature, seem to suggest one of the following. The first possibility is that this feature has spread into the Ngumpin-Yapa languages from the neighbouring non-Pama-Nyungan languages. Alternatively, if this feature used to be present in all (or most) Ngumpin-Yapa languages before, the fact that it has been preserved only in the most western and the most eastern Ngumpin-Yapa languages is then likely to be due to the contact with the non-Pama-Nyungan languages sharing the same feature. The second, retention

¹³ The following groups are not represented in my sample: I, K, O, P, R, S, V, NJ. Altogether, I have checked about a hundred “languages” (in Dixon’s terms). A given “language” is counted, when it has been possible to find information for at least one of its “dialects”. Sometimes, information for more than one “dialect” was available.

hypothesis appears to be the most plausible in view of the facts discussed in Section III.6.4.1.1.

6.3 Australian interrogative pronominals: some common functional and formal patterns

Before proceeding to the discussion of concrete languages and forms, let us first briefly examine some common functional (Sections III.6.3.1-III.6.3.2) and formal patterns (Sections III.6.3.3-III.6.3.4) of the interrogative pronominals in Australian languages. Most of them may be relevant for accounting for the lack of differentiation between ‘who?’ and ‘what?’.

6.3.1 *The interrogative/indefinite relation*

It is typical for Australian languages to use one and the same pro-form as an interrogative and an indefinite (cf. Dixon 2002:328; Mushin 1995), (2).

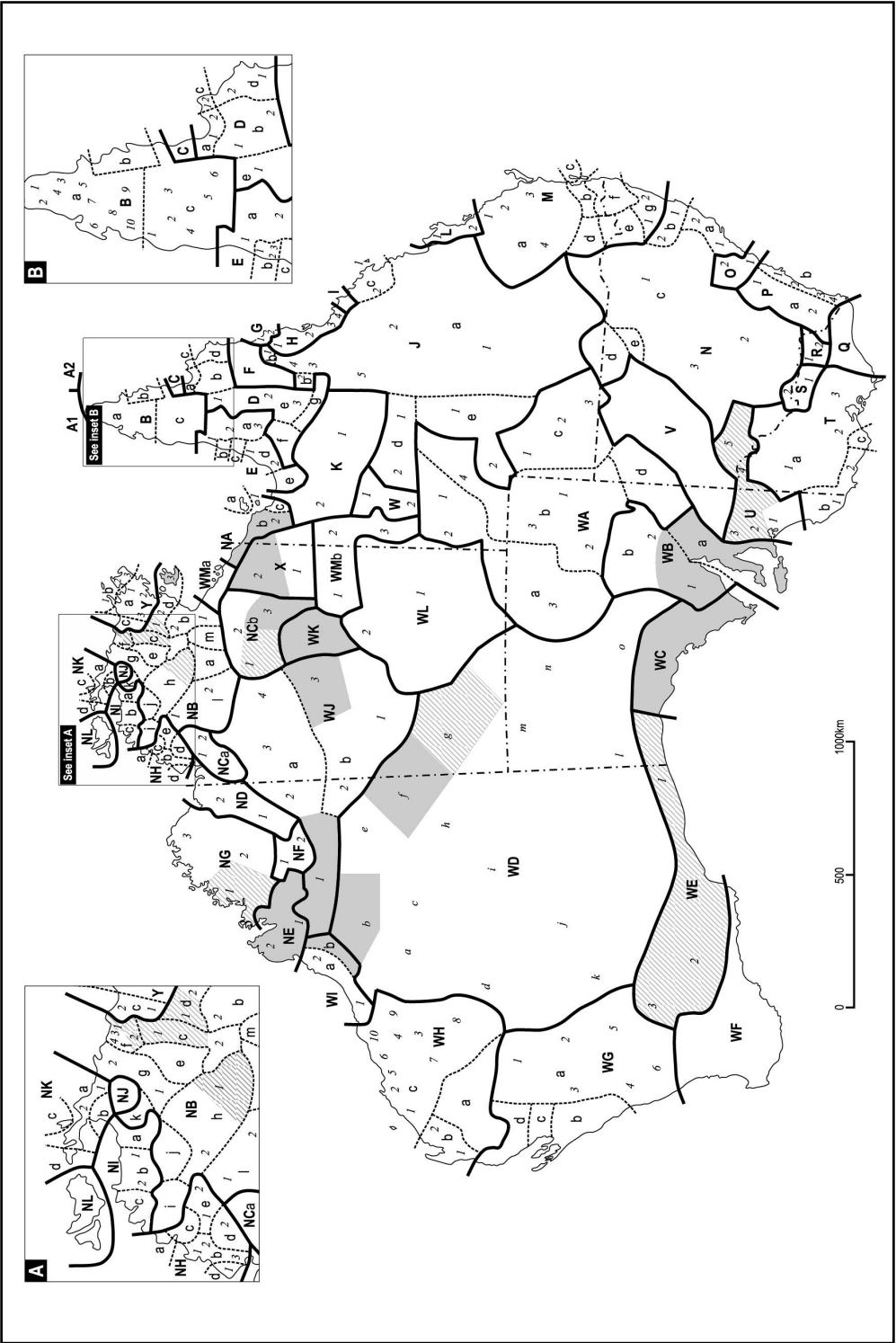
Yidiny (Pama-Nyungan; G2; Dixon 1977:182, 2002:328)

- (2) *wanyju* *walba* *yanggi:ny*
 who/someone.ERG rock.ABS split-PST
 ‘Someone must have cut the rock’ or ‘Who cut the rock?’, or ‘Someone must have cut the rock – who did it?’

In all probability, this feature of the Australian languages has also facilitated the semantic shift in many south-east Australian languages for a generic noun *minha* ‘(edible) animal’ to an indefinite/interrogative ‘something, what?’ described by Dixon (2002:334).

Dixon (2002:329) notes that “for most of the languages in which one form covers both ‘who’ and ‘what’, the available grammars do not give any additional indefinite sense”, with only “one known counter-example [...] Kayardild [(non-Pama-Nyungan, Tangkic; NAb1), where] *ngaaka* covers both ‘who’ and ‘what’ and also has the indefinite sense ‘someone’ and ‘something’”. However, as will be shown in Section III.6.4, “counter-examples” appear to be much more numerous than is assumed by Dixon. Thus, in at least 5 of the 16 “languages” that have proven to allow for a lack of differentiation between ‘who?’ and ‘what?’, the form used as the interrogative ‘who?, what?’ can also be used both as a human and a non-human indefinite pronominal, either on its own or with some additional morphology. The languages at issue are: Wirangu (Section III.6.4.1.1.2), Warumungu (Section III.6.4.1.1.5), Yulparija and Warlmanpa (Section III.6.4.1.1.6), and Warrwa (Section III.6.4.2.4). Kaurna (Section III.6.4.1.1.1), Wambaya (Section III.6.4.2.2.2), and Yawuru (which, in Dixon’s

Map 4. Australian languages that (may) have ‘who?, what?’ interrogatives



(1)	a. X2 WBa WBb1 WC WDb Wdf Wlb WJa1 WJb3 WK NAb1 NAb2 NBd3 NCb3 NE1 NE2	Garrwa (Eastern and Western) <i>Kaurna</i> , Kadli, Nantuwara, Ngadjuri, Narangka, Nukunu Parnkalla Wirangu Yulparija Pintupi Mangala Walmarjarri Warlmanpa Warumungu Kayardild Yukulta Anindilyakwa Wambaya Nyikina, Warrwa, Yawuru, Jukun Bardi, Nyulhyul, Jabirjabirr, Ngumbarl, Nimanburru	(1)	b. U2 U3 U4 U5 WDg WE1 WE2 WE3 NBc1 NBd1 NBf1 NBf2 NBh1 NCb1 NG1	Ngayawang Yuyu/Ngariket Keramin/Kureinji, etc. Yitha-Yitha, Dardi-Dardi Lurija Mirning Ngadjunmaya/Kalaaku Karlamay Rembarrnga Ngandi Burarra Gurr-goni Jawoyn Jingulu Worrorra

terms, is related to Warrwa as another dialect of the same language; Section III.6.4.2.4) may also belong here. The same applies to Jingulu (Section III.6.4.2.3), provided it really has a ‘who?, what?’ interrogative. At the same time, the situation with Dixon’s own counter-example, Kayardild, may be somewhat more complicated than Dixon’s wording may suggest (cf. Section III.6.4.2.5). For most other languages I do not have relevant information.

6.3.2 *The link between ‘where?’, ‘which [N]?, which one?’, ‘who?’ and ‘what?’*

According to Dixon (2002:327-328), in Australian languages there are often just three interrogative roots ‘who?’, ‘what?’ and ‘where?’, with “the other interrogatives [...] based on one of these three roots”. Dixon (2002:328) also notes that there rarely exist separate attributive interrogatives (“‘which’ is generally the same as ‘who’ in ‘which person’, the same as ‘what’ in ‘which thing’ and the same as ‘where’ in ‘which place’”). Mushin (1995:14) further notes that in Australian languages “a common property of the episteme PLACE [i.e., the word or the root meaning ‘where?’ and/or ‘somewhere’] is its use as an ‘epistemic determiner’ [i.e.] when used as a nominal modifier, it picks out set membership [which in English corresponds to] ‘which’”, that is also with non-locative nouns, (3, 4).

Muruwari (Pama-Nyungan; Nd; Oates 1988 via Mushin 1995:14)

- (3) a. *tirra yanta-a?*
 where-ABS has.gone-3SG
 ‘Where has she gone?’
 b. *tirra pathay tiya-rrī pu-ngka?*
 where-ABS father-ABS turn-REFL 2SG-GEN
 ‘Which one is your father?’

Mangarayi (non-Pama-Nyungan, Maran or Gunwinyguan; NBa; Merlan 1982:121)

- (4) a. *janan-gana ja-wula-nīja-n?*
 where-ABL 3-3PL-come-PRS
 ‘Where are they coming from?’
 b. *janan-gana Ø-malam-gana nya-may?*
 where-ABL M.ABL-man-ABL 2SG>3SG-take.PST
 ‘From which man did you take it?’

It is worth mentioning that besides the attributive function, as in ‘which [N]?’, in quite a few languages the root/word ‘where?’ may function pronominally, as ‘which one?’, (5b), (6b). Sometimes, the meaning ‘which one?’ is expressed by a combination of the locative interrogative with a demonstrative, (7b-c).

Ngiyambaa (Pama-Nyungan; Nc3; Donaldson 1980 via Mushin 1995:14)

- (5) a. *wanhdha-gu-ga: = na yana-nhi*
 INDF-DAT-IGNORATIVE=3ABS go-PST
 ‘He went somewhere, I don’t know where’
 b. *wanhdha-lu-wa: (miri-gu) dhingga: manundhiyi?*
 IPW-ERG-EXCLAMATIVE dog-ERG meat.ABS steal.PST
 ‘Which one (which dog) stole the meat?’

Mara (non-Pama-Nyungan, Maran; NBb1; Heath 1981:171-172)

- (6) a. *na-nguni na-na gariyi-mar?*
 M.SG-where M.SG-ART [M.SG]man-SG
 ‘Where is the man?’
 b. *na-nguni buribilipanji?*
 M.SG-where you.want.it
 ‘Which one do you want?’

Gun-djeihmi dialect of Bininj Gun-wok (non-Pama-Nyungan, Gunwinyguan; NBg1)

- (7) a. *ngayed yi-yo?*
 where 2-sleep
 ‘Where are you staying?’ (Evans 2003b:284)
 b. *ngayed na-be?*
 where M-DEM
 ‘Which one (masculine)?’ (Evans 2003b:286)
 c. *ngaye-ga na-be /al-de /an-de yi-jare?*
 where-LOC M-DEM F-DEM VEG-DEM 2-want
 ‘Which one (masculine/feminine/vegetable) do you want?’ (Evans 2003b:286)

Usually, it is not so difficult to demonstrate that ‘where?’ is the original function of the forms used as ‘which [N]?, which one?, where?’. First of all, a development from ‘which [N]?, which one?’ to the locative ‘where?’ would normally require some kind of additional locative morphology, whereas an extension from ‘where?’ to a selective ‘which [N]?, which one?’ may occur even without any additional morphology, as in (7b). Otherwise, some gender agreement morphology may be used, as in (8b). As such, an adverbial interrogative ‘where?’ is a less likely target for gender agreement (unless it functions as a predicate ‘(be) where?’) than a pronominal ‘which one?’ or an attributive ‘which [N]?’. Consider, in this respect, Ngankikurungkurr *kide* ‘where?’ (8a) that must be marked for gender agreement when used as a selective ‘which [N]?, which one?’ (8b-d).

Ngankikurungkurr/Nangikurrunggurr (non-Pama-Nyungan, Murrinh-Patha or Eastern Daly; NHd2)

- (8) a. *kide yerim fitat?*
 where 2SG.AUX.PRS put
 ‘Where did you put (it)?’ (Hoddinott & Kofod 1988:194)
- b. *kagu a-kide derrigirri yerim?*
 animal[G4] AG4-where like 2SG.AUX.PRS
 ‘Which is the meat you like?’ (Hoddinott & Kofod 1988:67)
- c. *a-kide wuddupun ta a-matyí?*
 AG4-where 3PL.PFV.AUX.PRS hit G4-kangaroo
 ‘Which is the kangaroo they killed?’ (Hoddinott & Kofod 1988:67)
- d. *wa-kide kagu wupun gat?*
 AG1-where animal[G4] 3SG.PFV.AUX.PRS catch
 ‘Which one (male human) caught the fish?’ (Hoddinott & Kofod 1988:67)

The roots of ‘which [N]?, which one?, where?’ interrogatives often contain frozen locative suffixes, which suggests the precedence of the locative meaning. Thus, the interrogatives similar to Ngiyambaa *wanhda-* (5), widespread in Pama-Nyungan languages, are believed to contain the frozen locative suffix *-dha* (Dixon 2002:165, 332, 334).

Given that ‘where?’ frequently expands to selective uses as ‘which [N]?, which one?’, it is hardly surprising to find also instances of its further extension to non-selective contexts, expectedly mostly as ‘who?’, rarely as ‘what?’ or both. For the non-Pama-Nyungan part of the continent consider, for instance, Wagiman *barrimiya* ‘who?’ and *barri* ‘where?’ (non-Pama-Nyungan, Wagiman-Wardaman; NB11; Wilson & Harvey 1999-2001),¹⁴ which may be cognate to Bininj Gun-wok *bale(?)* ‘where?, which one?, how?’ (all dialects except Gun-djeihmi; non-Pama-Nyungan, Gunwinyguan; NBg1; Evans 2003b:284-289). Another non-Pama-Nyungan language Ngalakan has AG-*werre* ‘who?’ and *werre-ka* ‘where?’ (Gunwinyguan; NBc2; Merlan 1983:77).¹⁵ Furthermore, Wambaya *gayini* ‘who? (= IPW.ABS.G1~G2)’, *gayina* ‘what? (= IPW.ABS.G4)’

¹⁴ Wilson & Harvey (1999-2001) spell *barrimiya* with a hyphen, viz. *barri-miya*, but they do not provide any translation for *-miya*. In all probability, the element *-miya* in *barrimiya* is cognate to the *-miya-* part of the nominals *manyimiyan* ‘this kind’ and *ganyimiyan* ‘that kind’.

¹⁵ Compare the Ngalakan locative/allative suffix *-ka?* ~ *-ga?* (in the allative sense also optionally as *-kaga?* ~ *-gaga?*) (Merlan 1983:40) and similar locative/allative suffixes in some other Gunwinyguan languages: Bininj Gun-wok *-ga* ~ *-ka(?)* ~ *-?ka?*, Dalabon and Rembarrnga ~ *-(?)ka?* (Evans 2003b:145).

(non-Pama-Nyungan, Mirndi group; NCb3; Nordlinger 1993:145-146) may be compared to Limilngan *gay*-AG ‘(be) where?’ (non-Pama-Nyungan; NIb; Harvey 2001:61-64), Burarra (AG-)*ga-ya* ‘where?’¹⁶ and AG-*an.ga-ya* ‘which one?, the one that is where?’¹⁷ (non-Pama-Nyungan, Maningrida; NBf1; Glasgow 1984:19, 43), and the Gun-djeihmi dialect of Bininj Gun-wok *ngayed* ‘where?, which one?’.

For the Pama-Nyungan part of the continent consider, for instance, Wagaya (or Wakaya) *winthi-nga* ‘who(M)?’, *winthi-nga-rr* ‘who(F)?’ (Pama-Nyungan; WMb1), which according to Breen (1974:3.3), “is clearly derived from *wiinthi* ‘where’ with the ‘after’ or ‘out of’ formative *-nga*”, as in *kirrii-nga* ‘spinster’ from *kirrii-rr* ‘single women’s camp, woman’, *wiiwa-nga* ‘stranger’ from *wiiwa(-r)l* ‘unknowing’ (possibly, lit.: ignorance-LOC) or *untumarni-nga* ‘bee’ from *untumarnt* ‘honey, bees’ nest’. Note that the final syllable *-thi* in *wiinthi* ‘where?’ is in all probability a frozen locative case marker. Consider also various *waar(r)*-like Pama-Nyungan interrogatives discussed by Dixon (2002:333), which mostly mean ‘who?’ and at least in the languages of the Eb group¹⁸ ‘where?’. The final *-r(r)*- is in all probability a reflex of some frozen locative suffix, comparable for instance to the one found in Nyangumarta *wanyja-rra* ‘where?’ vs. *wanyja-rni* ‘where?’ (Pama-Nyungan; WIa1; Sharp 2004:272). Another example can be found in Biri, where one root *nganhdha*- may be inflected for case to mean ‘who?’ or ‘where?’, or with some derivational morphology may mean ‘when?’ or ‘what cause?’ (Terrill 1998:27). The paradigm of *nganhdha* is reproduced in (9).

¹⁶ E.g., *yina an-ga-ya* ‘where is he?’ (lit.: Q M-“place”-REAL). Locative interrogatives based on the root *ga* can further be found, for instance, in Nunggubuyu *a-...-ga* ‘where?’ (non-Pama-Nyungan, Gunwinyguan; NBd2; Heath 1984:459-460) or Maranunggu *ka* ‘(be) where?’ (non-Pama-Nyungan, Western Daly; NHb1; Tryon 1970:71). Recall that for Australian languages the use of *k* or *g* is usually just a matter of spelling conventions.

¹⁷ The element *an-* is the frozen masculine gender marker.

¹⁸ Eb1 is Yir-Yoront (or Yirr-Yorront, Yirr-Thutjim) and Yirrk-Thangalkl (or Yirr(k)-Mel). Eb2 is Koko Bera (or Kok Kaber), Kok Peponk, Kok Wap, and Koko Beberam. Eb3 is Kok Thaw(a) (or Koko Petitj, Uw Inhal, Ogh Injigharr).

Biri (Pama-Nyungan; Ja2; Terrill 1998:27)

(9)	Transitive subject	‘who?’	<i>nganhdha-(ru)-nggu</i>
	Intransitive subject	‘who?’	<i>nganhdha-lu</i> or <i>nganhdha-ru</i>
	Object	‘whom?’	<i>nganhdha-lu-na</i> or <i>nganhdha-(ru)-na</i>
	Possessive	‘whose?’	<i>nganhdha-ngu</i>
	Dative	‘to/for whom?’	<i>nganhdha-(ru)-gu</i>
	Ablative	‘from whom?’	<i>nganhdha-dhamu</i>
	Semblative	‘like who?’	<i>nganhdha-(ru)-ngamu</i>
	Locative	‘where?’	<i>nganhdha-ru</i>
	Temporal	‘when?’	<i>nganhdha-ynbila</i>
	Cause	‘what cause?’	<i>nganhdha-mba</i>

Two things are worth emphasizing about the forms in (9). First, note that the “locative formative” *-ru* is also optional on most other case forms and the S-form ‘who?’ is even identical to the locative form ‘where?’. In other words, “it appears that the case forms are built upon the model of the locative case” (Terrill 1998:28). According to Terrill (1998:28), it is also “likely that all interrogative case forms have this optional *-ru*, but only some of these have been recorded”. Second, the root *nganhdha-* itself appears to be a frozen combination of the Biri interrogative *nganhi* ‘what?’ followed by an older locative suffix *-dha*, identical to the locative suffix mentioned above while discussing the Ngiyambaa interrogative *wanhdha-* ‘where?, which [N]?, which one?’ (5).¹⁹ In other words, in Biri the same evolution must have happened repeatedly.

In Duungidjawan (Pama-Nyungan; a “dialect” of Wakawaka, Ma4), Kite & Wurm (2004:66-67) report the consultant to use *wanyu-* ‘where (to, from)?, when?’ (10a) for ‘who?’ (10b), whereas the regular ‘who?’ is *ngan-* (10c).²⁰ Another Duungidjawan locative interrogative is *wanya/wanja* ‘where? (at, to)’ (10d).

Duungidjawan (Pama-Nyungan; “dialect” of Wakawaka, Ma4)

- (10) a. *wanyu-ngu mana ba-ye?*
 IPW-ABL DEM come-PRS
 ‘Where is (it) coming from?’ (Kite & Wurm 2004:66)

¹⁹ The interrogative *nganhi* itself seems to be lacking a locative form (cf. Terrill 1998:28).

²⁰ It should be mentioned that Kite & Wurm (2004:67) suggest that using *wanyu-* for ‘who?’ the informant “was mixing his languages” because in “the nearby Goreng-Goreng language [also known as Gureng Gureng] this is the form for ‘who’”. However, given the Biri facts just discussed, I would be somewhat reluctant to immediately resort to the hypothesis of mixing of languages. The fact that Goreng-Goreng (Ma2) is also rather closely related to Duungidjawan may suggest that Goreng-Goreng has simply already completed the evolution that in Duungidjawan was only on its way.

- b. *wanyu-nga binda-wu?*
 IPW-ACC send-FUT
 ‘Who will (we) send?’ (Kite & Wurm 2004:66)
- c. *ngan-du gana ba:ri-nyi?*
 who-ERG DEM bring-PFV
 ‘Who brought (the fish) here?’ (Kite & Wurm 2004:64)
- d. *wanja guwe ya-nyi?*
 where then go-PFV
 ‘Where did (they) go?’ (Kite & Wurm 2004:68)

The link from ‘where?’ to ‘what?’ appears to be much less common than the link from ‘where?’ to ‘who?’. In this respect, consider, for instance, Dixon’s (2002:332-334) overview of the *wanh*-based interrogatives in Australian languages. This comparison has shown that there are more than a hundred languages using this root for their ‘where?’ interrogatives, in most cases additionally marked with a (often frozen) locative suffix *-dha*, as in Ngiyambaa *wanhdha-* (5), although sometimes also without, as in Kuyani (Pama-Nyungan; WBb2; Hercus 1999:68) or Dhuwal (Pama-Nyungan; Ya; Morphy 1983) *wanha* ‘where?’ (cf. also Section III.6.4.1.1.1 on Kurna). Remarkably, Dixon also reports approximately 20 languages using the same root for their ‘who?’ interrogatives, but only approximately 4 using it for their ‘what?’ interrogatives. However, it should be pointed out that all the four also seem to base both their ‘who?’ and ‘where?’ interrogatives on the same (or very similar) root²¹ as well, as demonstrated in Table 1.

Table 1. The interrogatives ‘what?’, ‘who?’ and ‘where?’ in the four Australian languages that according to Dixon (2002:332) “have ‘what’ based on wanh-” (based on Breen 2003:441; Dixon 2002:330; Patz 2002:78-79)

	‘what?’	‘who?’	‘where?’
Girramay dialect of Dyirbal (Pama-Nyungan; H1)	<i>wanya</i>	<i>wanyunya</i> (ABS)	<i>wunyja-</i>
Yidiny (Pama-Nyungan; G2)	<i>wanyi</i>	<i>wanyju</i>	<i>wanyja</i>
Kuku-Yalanji (Pama-Nyungan; F)			
Yalanji dialect	<i>wanyu</i> (ABS)	<i>wanya</i> (ABS)	<i>wanja</i>
Nyungkul dialect	<i>wanyu</i> (ABS)	<i>wanju</i> (ABS)	<i>wanja</i>
Wanyi/Waanyi (non-Pama-Nyungan, Garrwan; X1)	<i>wanyi</i> (ABS)	<i>(w)inyjika</i> (ABS)	<i>winyja</i> (ABS)

²¹ Dixon (2002:332) himself suggests that forms with the first vowel *i* instead of *a*, such as *winya-* and the like, “are possibly cognate” to the root *wanh-*.

6.3.3 *Gender marking on interrogatives*

Many Australian languages, mostly the non-Pama-Nyungan ones, have gender. Typically, gender assignment appears to be largely semantic (cf. Dixon 2002:485-492; papers in Harvey & Reid 1997). Usually, all human nouns are fully divided between two genders, the feminine and the masculine. For non-humans gender assignment is generally less transparent, so that they may be assigned either to one of the non-human genders or to the genders containing human nouns. However, it should be pointed out that some semantic principles can still be established in many cases. For instance, bigger animals may be assigned to the masculine or the feminine gender according to their sex. Both for animates and inanimates, their mythological association, association with men's or women's activities, as well as various less direct associations may play an important role. Thus, the word for 'sun' is usually feminine, while the word for 'moon' is masculine. In Worrorra (non-Pama-Nyungan, Worrorran; NG1), the loan word for 'glass' is also feminine "because it shines like the sun, which is feminine", while "all tools used by the white man are described as masculine" (Love 2000:22). The Worrorra word for path, trail *kalumba*, which used to belong to one of the non-human genders, has become masculine after its semantics was expanded to refer to the modern road for vehicles, because it is "the Wo'rora men [who] have constructed a road for traffic" and "along this road come the loads of flour, tobacco and other goods, all of masculine gender in Wo'rora" (Love 2000:16). In Burarra (non-Pama-Nyungan, Maningrida; NBf1), the masculine gender besides human males also includes "many animals, the moon and metal objects", while the feminine besides human females also includes some animals and the sun (Glasgow 1984:7).

Many languages also mark gender on their interrogative pronominals and quite a few on the locative interrogative 'where?' (cf. Dixon 2002:478-479; cf. also Section III.6.3.2). Some languages, such as Ngandi (Gunwinyguan; NBd1) mentioned in Section III.6.2, build their interrogative pronominals on a single root and use different gender markers to distinguish the human interrogative 'who?' from the non-human interrogative 'what?'. In the latter case, the peculiarities of the semantics of genders and agreement patterns in a given language may (i) influence the choice between 'who?' and 'what?' in cases of non-prototypical combinations of values, (ii) result in some typologically very unusual uses of interrogatives in "no conjecture"-contexts or (iii) result in a partial lack of differentiation between 'who?' and 'what?'. Let us consider some of such cases, starting with non-prototypical combinations of values.

The kind of non-prototypical combination of values I will discuss is represented by ANIMATE-questions involving the use of 'who?', viz. [animate thing + classification (+ common noun)] (cf. Section II.4). The use of 'who?' here is due to the fact the gender(s) containing human nouns also often contain

animals, so that, for instance, a given language has a masculine *animate* gender rather than a masculine *human* gender. As a result the same interrogative will be used both in questions about persons and animals. For instance, in Wambaya, which has four genders, of which G1 can be labelled as masculine and G2 as feminine, the masculine interrogative pronominal *gayini* is the default form in questions about humans and thus the closest equivalent of ‘who?’ (non-Pama-Nyungan, Mirndi group; NCb3; cf. Section III.6.4.2.2.2), as in (11).

Wambaya (non-Pama-Nyungan, Mirndi group; NCb3)

- (11) *gayini* *g-a* *yarru* *ginkanyi nanganangali?*
 IPW.M.SG.ABS 3SG-PST go[NON<FUT>] this.way sneak.away.NON<FUT>
 ‘Who was it that snuck off this way?’ (Nordlinger 1993, example 7-89)

However, the masculine interrogative must also be used in questions about animals when the word *yangaji* ‘animal, meat’ is used, as in (12), because the latter word is masculine.²²

Wambaya

- (12) *gayini* *ini* *yangaji?*
 IPW.M.SG.ABS this.M.SG.ABS animal.M.SG.ABS
 ‘What animal is this?’ (Nordlinger 1998:3.5)

Given that all animals in Wambaya belong either to the masculine or the feminine gender (cf. Nordlinger 1993:4.2.1), the two genders are best viewed as the masculine animate and the feminine animate genders respectively.

When the use of ‘who?’ in ANIMATE-questions is not mentioned explicitly in a given source, it may still be reasonable to assume its existence in a language about which it is known that animals are assigned to the human gender (or genders) and where an interrogative pronominal root, as such indifferent to the human vs. non-human distinction, is marked for gender. At least the following languages are likely to use ‘who?’ in ANIMATE-questions in the same way as discussed for Wambaya above. To begin with, recall Burarra and Worrorra mentioned in the beginning of this section and discussed in somewhat more detail in Section III.6.4.2.2.1. Both Burarra and Worrorra have two human genders, masculine and feminine, next to two exclusively non-human genders. The meanings ‘who?’ and ‘what?’ are expressed in Burarra with one and in Worrorra with two interrogative pronominal stems marked by the appropriate gender marker. The interrogative pronominal stem in Burarra is *-(yi)nga*

²² Apparently, if the G4 word *gunju*, a “less commonly used synonym” of *yangaji* (Nordlinger 1993:4.2.1), were used instead, the interrogative would be of G4 as well. The word *gunju* is likely to be a loan from Nungali, which in all probability has been assigned to G4 “on the basis of its form” (Nordlinger 1993:4.2.1, fn. 59).

(Glasgow 1984:42). Worrorra has *angku-* for the masculine and the feminine and *angu-* for the two non-human genders (Clendon 1994 via Dixon 2002:476; Love 2000:16-17). However, the two genders to which humans are assigned, the masculine and the feminine, also contain most other animates, e.g. in Worrorra “most of the birds and flies are feminine” (Love 2000:22). Gurr-goni/Guragone (non-Pama-Nyungan, Maningrida; NBf2) and, to a lesser extent, Ngandi (non-Pama-Nyungan, Gunwinyguan; NBd1) appear to be similar to Burarra. Jawoyn (non-Pama-Nyungan, Gunwinyguan; NBh1) has a single interrogative pronominal root marked for gender and the masculine gender appears to be also the “animate” gender (cf. Dixon 2002:328, 479, 506-508).

In some languages of northern Australia, the use of gender marking on an interrogative pronominal stem to distinguish ‘who?’ from ‘what?’ may result in a very peculiar use of non-selective interrogative pronominals in “no conjecture”-contexts, which does not appear to occur elsewhere in the languages of the world. In at least two Australian languages, Wambaya, already mentioned above, and Gurr-goni/Guragone, the masculine form of the interrogative pronominal, which is the default form in questions about humans and thus the closest equivalent of ‘who?’, is also used “when the referent is unknown, such that it is not possible to determine the gender” (Nordlinger 1993:4.7.1 about Wambaya; cf. also Section III.6.4.2.2.2), (13-14).

Wambaya (non-Pama-Nyungan, Mirndi group; NCb3)

- (13) *gayini* *irri-n* *ngannga?*
 IPW.M.ACC 3PL.A.[NON<PST>]-PROG bark.at[NON<FUT>]
 ‘What are they barking at?’ (Nordlinger 1993, example 4-228)

Gurr-goni (non-Pama-Nyungan, Maningrida; NBf2)

- (14) *a-nyi* *nyi-na-ni??*
 M-IPW 2MIN>3MIN-see-PRECONTEMPORARY
 ‘What did you see?’ (Green 1995:64-65 via Dixon 2002:489)

This is not a lack of differentiation between ‘who?’ and ‘what?’ because, apparently, the masculine interrogative pronominal will not be used when the speaker is clearly aware that the referent is non-human (and non-masculine). Note that unlike Wambaya and Gurr-goni, English would rather use the non-human interrogative ‘what?’ in such a context.

The use of the masculine interrogative pronominal in “no conjecture”-contexts in Wambaya and Gurr-goni is due to the default status of the masculine agreement pattern in these two languages.²³ The default use of the masculine

²³ Jawoyn (non-Pama-Nyungan, Gunwinyguan; NBh1) might be another language using its interrogative pronominals in the same way as Wambaya and Gurr-goni in “no conjecture”-contexts. It has all the necessary “ingredients”: a general interrogative pronominal root marked

agreement pattern on certain targets and in certain contexts appears to be quite common among the non-Pama-Nyungan languages (cf. e.g. Harvey 2001:46 on the non-Pama-Nyungan languages of the western Top End). Evans (1997) calls this “superclassing” in his description of gender and agreement in Bininj Gunwok dialect cluster.

Finally, in some languages the use of an interrogative pronominal stem, as such indifferent to the human vs. non-human distinction, with different gender marking appears to lead to a lack of differentiation between ‘who?’ and ‘what?’ in certain contexts. The reason is that in some of these languages, besides animals, the human gender(s) may include a good deal of inanimates. Sometimes, inanimates appear to be assigned to the human gender(s) following some clear semantic principles, as in Burarra and Worrorra (cf. above and Section III.6.4.2.2.1). Sometimes, the assignment may be less transparent, as in Wambaya (Section III.6.4.2.2.2). Other languages that might belong here as well are Ngandi (non-Pama-Nyungan, Gunwinyguan; NBd1), Gurr-goni/Guragone (non-Pama-Nyungan, Maningrida; NBf2) and Jawoyn (non-Pama-Nyungan, Gunwinyguan; NBh1). At least in Ngandi and Jawoyn, the masculine and feminine genders have proven to contain some inanimates (cf. Harvey 1997a:50-51, 55-56, 1997b:150-151). However, I have not been able to check whether the masculine and feminine forms of the interrogative pronominal can also be used in questions about inanimates.

6.3.4 *Shortening and augmentation*

Two opposite processes, shortening and augmentation of the interrogatives, appear to be common in Australian languages. Their importance can hardly be underestimated, since shortening and augmentation of the interrogatives, especially when repeated a few times, may change two related forms to such an extent that their common origin will be barely recognizable. For instance, when one compares just the Pitta-Pitta and Yankunytjatjara forms of the interrogative ‘where?’ in (15) below, *winhtha-* and *yaaltyi(-)* respectively, they would appear to be completely unrelated. Their common origin becomes apparent only when forms from other languages are also taken into consideration. However, a similar approach may be much more difficult to apply in the case of the non-Pama-Nyungan languages, primarily because of the considerable depth of the genetic relations between most languages and because in many cases we lack the certainty about which languages are related and to what degree (Section III.6.1). In what follows, let us consider the two processes in more detail, starting with shortening and then proceeding to augmentation.

for gender (Capell 1942:385-386; Heath 1978) and the masculine agreement pattern that, by and large, appears to have the default status (cf. Harvey 1997b:150-151).

Generally speaking, a given form may be shortened in two different ways. First, it may be shortened due to some more or less natural phonological process, like vowel length reduction, loss of phonologically less prominent segments, coalescence, etc. In the Australian context, one of the most prominent changes of this kind is probably the so-called “initial dropping”, i.e. the dropping of the initial consonants and vowels due to the peculiarities of the stress phonetics in the languages of Australia (cf. Dixon 2002:589-591). Another kind of change which leads to shortening of a given form but is not directly attributable to any regular phonological process in the language at issue may be called clipping. When clipping is applied to a word, normally whole syllables are dropped, not just phonemes, as in *advertisement* when it becomes *ad*. Parts of a more or less fixed expression may be clipped out as well, e.g. *pub* from *public house*.

In his overview of the interrogatives in Australian languages, Dixon (2002:329-330) notes that “there appears to be a tendency to shorten interrogative forms [through] omit[ting] the initial syllable of an interrogative form”, even in languages “which do not manifest any general initial-dropping changes”, such as languages “where stress generally goes on the initial syllable”. For instance, Dixon compares the Wangka-yutjuru (Pama-Nyungan; WAa2) interrogatives *thanha-* ‘where?’ and *thilampara* ‘how?’ to the interrogatives *winhtha-* ‘where?’ and *withila* ‘how?’ in the closely related language Pitta-Pitta (WAa1) and suggests that “it is likely that Wangka-yutjuru has simply omitted the initial syllable, *wi-*, from each form”. The Wangka-yutjuru and Pitta-Pitta ‘where?’ interrogatives can be further compared to similar interrogatives in some of the genetically and/or geographically close Pama-Nyungan languages, as summarized in (15).

- (15) *The interrogative ‘where?’ in some genetically and/or geographically close Pama-Nyungan languages (based on Dixon 2002:329; Goddard 1994:249-250; Hercus 1999:65, 68; Wilkins 1989:3.8.1)*

	‘where?’
Pitta-Pitta (WAa1)	<i>winhtha-</i>
Wangka-yutjuru (WAa2)	<i>thanha-</i>
Mparntwe Arrernte (WL1)	<i>nthenhe</i>
Wirangu (WC)	<i>indhā, (in)dhala(-)</i>
Arabana (WAa3)	<i>indya-, intya-</i>
Wangganguru/Wangkangurru (WAa3)	<i>intyali</i>
Kukata/Kukarta (WDo)	<i>nyaal</i>
Yankunytjatjara (WDn)	<i>yaaltyi(-)</i>

The elements *-dha*, *-tha*, *-ja*, and *-tya*, as well as *-l(a)*, in the forms in (15) are frozen locative suffixes (cf. Section III.6.3.2). According to Goddard (1994:250), the part *-tyi* in Yankunytjatjara *yaaltyi(-)* goes back to *alatyī* ‘like this’, “the

deictic used in conjunction with acts of demonstration”. The elements *-nh(a)*, *-nhe* and probably also *-n-*, may go back to a demonstrative or the final syllable thereof. Thus, Mparntwe Arrernte has *nhenhe* ‘this [N], this one, here’ and *yanhe* ‘that [N], that one, there’, as well as *ngwenhe* ‘who?’ and *iwenhe* ‘what?’ (Wilkins 1989:3.6.1, 3.8.1). Ultimately, this final syllable of demonstratives should probably be brought back to a specificity/focus marker, similar to the “specific referent marker” *-nha* of Panyjima, for instance (cf. below).

Another example of apparent clipping, mentioned by Dixon and reproduced here in (16) is that of Djabugay (Pama-Nyungan; G1) as compared to the closely related Yidiny (Pama-Nyungan; G2).

(16)	Yidiny	Djabugay
‘vegetable food’	<i>mayi</i>	<i>ma:</i>
‘who?’	<i>wanyju</i>	<i>ju:</i>
‘what?’	<i>wanyi</i>	<i>nyi:</i>
‘where?’	<i>wanyja</i>	<i>ja:</i>

It is remarkable that “for the lexeme ‘vegetable food’ it is the final syllable that has been omitted, but for the three interrogatives it is the initial syllable (in each instance the vowel has been lengthened since all monosyllabic words in Djabugay must involve a long vowel)” (2002:330).

It has already become clear from some of the examples cited above that while interrogatives may often be shortened at their left edge, they are also often augmented at their right edge (in some prefixing languages also at the left). The elements that apparently tend to become frozen on the right edge of the interrogatives most often are case markers and demonstratives (cf. examples (9) and (15) above).²⁴ Other elements that in principle have good chances to be used for augmenting the interrogatives are various emphasis markers, (meaningless) increments and in the case of gender languages probably also gender markers.

For instance, according to Donaldson (1980 via Mushin 1995:4) in Ngiyambaa (Pama-Nyungan; Nc3) the interrogative-cum-indefinite pro-forms must occur with one of the two knowledge clitics, =*wa:* ‘exclamative’, which marks interrogativity (5b), or =*ga:* ‘ignorative’, which marks indefiniteness (5a).

²⁴ The interrogatives of Limilngan (Nlb1) may be adduced as an example from the non-Pama-Nyungan area. In Limilngan, the interrogative pronominals ‘who?’ and ‘what?’ are based on the same root *inyi-* followed by a gender marker and by a frozen “distal suffix” *-k*; the same is true for the interrogative ‘where?’ which is just based on a different root, *gay-* (Harvey 2001:59). Compare, for instance, the human gender G1 forms of the demonstratives and interrogatives: *da-wi-k* ~ *ja-wi-k* ‘that, there’, *da-wi-n* ~ *ja-wi-n* ‘this, here’, *inyi-wi-k* ‘who?’, *gay-wi-k* ‘where?’, but no **inyi-wi-n* or **gay-wi-n*.

Pintupi optionally uses the suffix *-pa* on interrogatives (Pama-Nyungan; Wdf; Hansen & Hansen 1978:216). In Mangarayi (non-Pama-Nyungan, Maran or Gunwinyguan; NBa; Merlan 1982:6), “WH-words tend to be the first element in the clause” and “when clause-initial, they tend to be cliticized with *-bayi* or its reduced form *-ba* (termed ‘focus’ clitic [...])”. In Wirangu (Pama-Nyungan; WC), the interrogative *ngana* ‘who?, what?’ is also found to be used with “the suffix *-(g)ardu*, *-(g)arda*” (17), which is otherwise “an emphatic suffix” occurring “only with a limited group of adjectives, [as well as] deictic pronouns” (Hercus 1999:29, 91).

Wirangu (Pama-Nyungan; WC; Hercus 1999:91)

- (17) *ngana-arda bala-arda?*
 IPW-FOC this-FOC
 ‘What is this?’

In Panyjima/Panytyima (Pama-Nyungan; WHc3), the interrogative *tharni* ‘where?’ may be marked with the so-called “specific referent marker” *-nha* (the accusative form is *-ngu*) to inquire about “an unknown but specific location” (Dench 1991:146-147, 165-166),²⁵ as in (18b) vs. (18a).

Panyjima (Pama-Nyungan; WHc3; Dench 1991:165-166)

- (18) a. *tharni-wali yana-ku-rru?*
 where-ALL go-PRS-NOW
 ‘Where are you going?’
 b. *tharni-nha-wali yapal-purlu-nta?*
 where-SPECIFIC-ALL downriver-face-INTENTIONAL
 ‘Where exactly are you going, downriver?’

In certain languages, meaningless increments, in the form of a syllable or a single phoneme, must be added to fulfil various phonotactic constraints, such as minimality constraints or the ban on words/roots ending in consonants, etc. (cf. Dixon 2002:65, 647-650). For instance, in many Pama-Nyungan languages in the western and central part of the continent a syllable *-ba/-ma* is added to either word or root ending in a consonant, as in Mantharta *nguwana^{ma}* ‘sleep’ and *yuwalba* ‘wind’ (Austin 1980:50 via Dixon 2002:647). In Ngayawang (Pama-Nyungan; U2), the increment used “to avoid a final consonant” is *-ko* (Blake 2003:17). In Lardil (non-Pama-Nyungan, Tangkic; NAa), only monosyllabic

²⁵ The “specific” forms of the interrogatives *ngana* ‘who? (NOM)’ and *ngananha* ‘what? (NOM)’ are somewhat irregular, *nganapukul(pa)* and *ngapinha* respectively (Dench 1991:164-165). Note, however, that the accusative of *ngana* is *ngana-ngu*, i.e. it is marked with the specific marker. Furthermore, the final *-nha* in the forms for ‘what?’ may represent the frozen maker *-nha*.

roots have a non-zero nominative, marked as *-a* if the root ends in a vowel and *-ka*, *-ta* or *-rta* if the root ends in a consonant (the allomorph has the same place of articulation as the final consonant), for pronouns the augment is *-(k)i* (Klokeid 1976:54-56, 115-116). As has already been mentioned, vowel lengthening in the (shortened) forms of the Djabugay interrogatives in (16) is also due to the constraint on CV words. According to (Breen 2003:435), in two Garrwan languages, Wanyi (X1) and Eastern Garrwa (X2), a “suffix *-n* is frequently attached to vowel-final noun stems, both in citation form and in sentences”, but “no function is known for this”. However, in Western Garrwa (X2) this *-n* occurs only “on the pronoun form of some suffixes”, e.g. the ‘origin’ suffix on demonstratives is *-nmungkuji*, while on other nominals it is simply *-mungkuji* (Breen 2003:435).

Gender markers may become frozen both on targets and controllers due to various reasons, such as shift from prefixing to suffixing in gender marking, gender neutralization, etc. Some examples of possible cases of loss of the morpheme status of gender markers can be found, for instance, in Dixon (2002:475-476, 506-508) or Harvey, Green & Nordlinger (2006). An interesting example can be found in the Kune dialect of Bininj Gun-wok (non-Pama-Nyungan, Gunwinyguan; NBg1). Kune has fully neutralized gender agreement on demonstratives and adjectives in favour of the original masculine agreement marker *na-*, so that now all demonstratives (always) and adjectives (almost always)²⁶ begin in *na-*, even though the original gender markers on the nouns have been preserved intact (cf. Evans 2003b:181-182, 308-309). In Mara (non-Pama-Nyungan, Maran; NBb1; Heath 1981:171-174), the interrogative root *-nguni* ‘where?’ when marked by a dual or a plural prefix shows up as *-ɲanguni*. The latter form is identical to the feminine singular form of ‘where?’ *ɲa-nguni* and almost identical to its neuter form *n-ga-nguni*. Furthermore, note that there exist two other locative interrogatives ‘where?’, *gangu* and *ɲangu*, which are not marked for gender. In Burarra (non-Pama-Nyungan, Maningrida; NBf1), the interrogative ‘which one?, the one that is where?’ AG-*an.ga-ya* is derived from the locative interrogative construction *yina* (AG)-*ga-ya* and *an-* in ‘which one?’ is the frozen masculine gender marker *an-* (Glasgow 1984:43; cf. also Section III.6.3.2).

²⁶ Evans (2003b:182) notes that the feminine agreement pattern may still be encountered occasionally and the initial *na-* of the adjectives is also dropped when the “adjective is followed by a verbalizing suffix”.

6.4 A closer look at the Australian languages with ‘who?, what?’ interrogatives

In what follows, the discussion will be organized along genetic lines, viz. Pama-Nyungan (Section III.6.4.1) vs. non-Pama-Nyungan (Section III.6.4.2). This division is justified by the fact that Pama-Nyungan languages appear to be much more homogeneous in the forms and origins of the presumed ‘who?, what?’ interrogatives as compared to the non-Pama-Nyungan languages.

6.4.1 *Pama-Nyungan languages*

The Pama-Nyungan languages with an apparent lack of differentiation between ‘who?’ and ‘what?’ fall into two groups and will be examined correspondingly. The first and by far the biggest group, discussed in Section III.6.4.1.1, comprises languages where the ‘who?, what?’ interrogatives appear to go back to an earlier form *ngana* ‘who?, what?’. The second group, Section III.6.4.1.2, includes two possible cases of the lack of differentiation between ‘who?’ and ‘what?’ that do no fit in the first group and for which the data have been too meagre to allow for any definite conclusions.

6.4.1.1 The interrogatives *ngana* > *ngana-nha* > *nhaa* and the like

Several Pama-Nyungan languages in the western and central part of the continent²⁷ use a single interrogative as both ‘who?’ and ‘what?’. These ‘who?, what?’ interrogatives appear to go back to an earlier form *ngana* ‘who?, what?’, which is likely to go back at least to Proto Pama-Nyungan. In some languages, *ngana* has been later incremented with a specificity/focus marker *-nha/-nya*. Subsequently, in many languages the incremented and the base form have become differentiated as ‘what?’ and ‘who?’ respectively. An important factor in the latter process must have been the common tendency of the marker *-nha/-nya* to be reinterpreted as an accusative/absolutive case marker for proper names and pronominals. In many languages, especially in those where the base form and the incremented form have become semantically differentiated along the human vs. non-human lines, the incremented form has been clipped to *nhaa/nha-* and the like. In some languages, the latter clipped form has been subsequently augmented with other morphemes.

Geographically, the Pama-Nyungan languages discussed in what follows fall into two groups, the first one stretches along the coast of the Great Australian Bight, while the languages of the second group are found in the northern fringes

²⁷ By the western and central parts of the continent I understand the territory covered by the states of Western and South Australia and most parts of the Northern Territory.

of the Pama-Nyungan area. The languages in the south are either moribund or already extinct. The situation is somewhat better in the north. In what follows, I will first discuss the southern languages: Kurna and Parnkalla in Section III.6.4.1.1.1, Wirangu in Section III.6.4.1.1.2, Mirning, Ngadjunmaya and Karlamay in Section III.6.4.1.1.3. Then the languages further to the north will be examined: Pintupi in Section III.6.4.1.1.4, Warumungu in Section III.6.4.1.1.5, Yulparija, Mangala, Walmajarri and Warlmanpa in Section III.6.4.1.1.6, and Mangala in Section III.6.4.1.1.7.

6.4.1.1.1 Kurna and Parnkalla

Kurna (WBa) and Parnkalla/Banggarla (WBb1) are two Pama-Nyungan languages of the so-called Thura-Yura group. Kurna, as described by Teichelmann & Schürmann (1840) used to be spoken “in and for some distance around Adelaide”. Parnkalla, as described by Schürmann (1844) used to be spoken a bit further to the north, on the Eyre Peninsula. According to Hercus (1999:91), both Parnkalla and Wirangu have “just one form, ‘nganna’, for ‘who’ and ‘what’”, while “the northern Thura-Yura languages Adynyamathanha and Kuyani [both WBb2] have two separate stems”. The information on Adynyamathanha is supported by Schebeck (1974:13). I have not been able to check Schürmann’s (1844) description of Parnkalla and therefore I have to rely on the aforementioned statement by Hercus. For Kurna I have been able to consult the original source, Teichelmann & Schürmann (1840). According to this source, Kurna uses an interrogative pronominal *nganna* ‘who?, what?’. Its case-number paradigm is reproduced in Table 2. Besides, there may be another unrelated ‘who?, what?’ interrogative *wada(i)(n)na*, which will be discussed in the end of this section.

Table 2. The case-number paradigm of the Kurna interrogative pronominal *nganna* (adapted from Teichelmann & Schürmann 1840:I.9)

	SG	DU	PL
ABS	<i>nganna</i>	<i>ngandourla</i>	<i>ngandoanna</i>
DAT	<i>nganna</i>	<i>ngandurla</i>	<i>ngandoanna</i>
GEN	<i>ngangko</i>	<i>ngandourlakko</i>	<i>ngandoannako</i>
ERG	<i>ngannarlo</i> ‘what?’ <i>ngando</i> ‘who?’	—	—

However, the data on Kurna in Teichelmann & Schürmann (1840) allow for at least two different interpretations. First of all, note that Teichelmann &

Schürmann (1840:I.10) point out that the “irregular [ergative]”²⁸ form [*ngando*] is applied when the cause of an action or effect is supposed to be a person or being” (19a), while the regular form *ngannarlo* is used “when the cause is either unknown or an instrument” (19b). “For the dual and plural of [the ergative] case no termination is known” (Teichelmann & Schürmann 1840:I.6).

Kaurna (Pama-Nyungan; WBa; Teichelmann & Schürmann 1840:I.10)

- (19) a. *ninna* *ngannarlo* *minkarni?*
 2SG.ABS IPW.ERG has.wounded
 ‘By what have you been wounded?’
- b. *ngando* *aityo* *mudlinna* *metti?*
 IPW.ERG 1SG.POSS.ABS implement.ABS.PL has.taken.away
 ‘Who has taken away my implements?’

This seems to suggest that although the distinction between ‘who?’ and ‘what?’ in Kaurna is indeed largely inexistent, the two meanings are distinguished in the ergative singular.

Alternatively, provided we analyze the two ergative suffixes in a slightly different way, Kaurna may also be considered as a language with a complete lack of differentiation between ‘who?’ and ‘what?’. The following facts may be adduced here. To begin with, note that Dixon (2002:157-159; largely basing himself on Sands 1996), reports that generally, in Pama-Nyungan languages two allomorphs of the ergative suffix can be found, the first one, *-lu*, originally in all probability restricted to proper nouns, and the other *-dhu*, used on common nouns.²⁹ Remarkably, in Adynyamathanha (WBb2), rather closely related to Kaurna, we find a somewhat similar differentiation. Thus, according to Schebeck (1974:3-4), Adynyamathanha has the ergative suffix *-lu*, which “combines with proper nouns, but only with certain ‘common nouns’ and pronouns”, and the suffix *-nga*, which “the great majority of nouns [also including proper nouns] take [...] to mark the agentive [i.e., the ergative], and all nouns can take this suffix in its other main functions which are” the locative, comitative, temporal and ‘in what concerns [N]’. It may be hypothesized that in Kaurna the situation was somewhat similar. Consequently, the human meaning of the “irregular” form *ngando* and the non-human meaning of the “regular” form *ngannarlo* would be due to the semantics of the respective case suffixes, *-do* and *-rlo*.

Unlike in many other Australian languages, in Kaurna the forms used as interrogative pronominals do not seem to be used as indefinites. However, the

²⁸ Teichelmann & Schürmann call the ergative “the active or ablative case” (1840:I.6).

²⁹ There is also the third ergative allomorph, *-nggu*, which is somewhat less common and usually also seems to be restricted to common nouns in the languages where this parameter is relevant (Dixon 2002:159-161).

Kaurna indefinite pronominals *ngapidlo/ngapillo* ‘somebody, something’ and *ngapidlurlo* ‘somebody (“was the agent”)’ (Teichelmann & Schürmann 1840:I.13, II.31) are clearly related to the interrogatives. Since both *-dlo* and *-(u)rlo* are regular ergative suffixes, it is not clear whether the second form is the ergative of the first one or the two are just different indefinites.

Let us now discuss in more detail the form of the Kaurna interrogative stem *nganna* itself. Its final segment *-na* deserves particular attention. I believe that this is the same element as the Adynyamathanha suffix *-nha*, so that (at least etymologically) Teichelmann & Schürmann’s *nganna* is likely to be really *ngananha* ‘who?, what?’, similar to Wirangu *ngananha* ‘who?, what?’ (WC; cf. Section III.6.4.1.1.2), Panyjima *ngananha* ‘what?’ (WHc3; Dench 1991:164) or Pitjantjatjara *ngana-nya* ‘who-ABS?’ (WDm; Eckert & Hudson 1988:119).

In Adynyamathanha, although “the functions of the suffix *-nha* are doubtlessly the most difficult to define”, a few common uses can be recognized (Schebeck 1974:4). For instance, it can be used with the absolutive function (which otherwise is normally unmarked), and “the use of this suffix is practically the rule with place names and proper names” (Schebeck 1974:4-6). This suffix also occurs on bound demonstrative roots, e.g. *ngu-nha* ‘that over there’ vs. *ngu-tla* ‘over there somewhere, over there you know’, and may be “sometimes [used] to confer the value of a substantive to a demonstrative pronoun”, e.g. *i-nha* ‘this’ vs. *i-nha-nha* ‘this one’ (Schebeck 1974:11-12). The element *-nha* also appears on interrogatives (and words like ‘(an)other (one)’). Thus, among other things, we find (i) *wa-nha(-)* ‘where?’, “consist[ing] of the interrogative *-wa* and the supporting element *-nha*”, (ii) *nganha-* ‘who?’, (iii) *nhangata-* ‘what?’, which “always has the suffix *-nha* in the ‘nominative’ [i.e., absolutive]” (Schebeck 1974:13-14). Remarkably, the interrogative ‘what?’ drops *-nha* “before some suffixes (e.g. *nhangata-ju* ‘what for? why?’), but not before others (e.g. *nhangata-nha-nga* ‘in something’ [where *-nga* is the ergative/locative suffix mentioned above])” (Schebeck 1974:13; emphasis added). This is strongly reminiscent of what one finds in Kaurna, where the final segment “*-na*” of “*nganna*” is dropped before the genitive *-ko* and the ergative *-do*, but is maintained in the dative and before the ergative *-rlo*.

The functions that *-nha* normally fulfils in Adynyamathanha strongly suggest that ultimately it should probably be brought back to the same specific marker that frequently becomes frozen on the interrogative pronominals of the west and central Australian Pama-Nyungan languages. This is the same marker as found, for instance, on the Mparntwe Arrernte interrogatives and demonstratives or in Panyjima as the “specific referent marker” *-nha* (cf. Section III.6.3.4), as well as in Warumungu as the “emphatic and deictic” suffix *-ya* (cf. Section III.6.4.1.1.5).

By way of conclusion, let us briefly consider yet another interrogative that Teichelmann & Schürmann gloss as ‘who?, which one?, what?’, *wādaina* ~

wadana ~ *wādanna* (1840:II.50, 1840:I.10). This interrogative seems to be based on the interrogative *wāda* ~ *wada* ‘where?’, “used only when an other person has stated something referring to a locality” (1840:II.50), and the demonstrative *inna* ‘this one’ (1840:II.6).³⁰ Thus, its literal meaning must have been something like ‘where one?’, i.e. ‘which one?’. Since no examples are provided, it is difficult to say whether this interrogative can really be used in non-selective questions as both ‘who?’ and ‘what?’, even though semantically this kind of development would be quite plausible (cf. a similar reservation concerning Lurija *yaal* in Section III.6.4.1.2).

6.4.1.1.2 Wirangu

Wirangu (WC), as described by Hercus (1999), is a Pama-Nyungan language spoken immediately to the west of the Thura-Yura languages Parnkalla (WBb1), Adynyamathanha (WBb2) and Kaurna (WBa), discussed in Section III.6.4.1.1.1. The position of Wirangu within Pama-Nyungan has been the subject of some debate. Simpson & Hercus (1996) argue that Wirangu is an “outlier of the Thura-Yura group of languages” (cited via Hercus 1999:10; see also Hercus 1999:1, 8-15 for a discussion and further references).

³⁰ The sound value of the grapheme *ā* is not explained by Teichelmann & Schürmann. Besides, it appears to be used somewhat inconsistently: being rare as such, it is also often replaced by the simple *a* (“sounds as the same letter in *harp*, *hard*”, 1840:I.2) and in few cases by *ã* (“sounds as the same letter in *wall*, *ball*”, 1840:I.2), although the variation in the opposite direction, i.e. from *a* or *ā* to *ã*, seems to be very rare. Given that (i) the simplest word for ‘where?’ in Kaurna is *wā* ~ *wa* (1840:II.50), (ii) in the not so distantly related Adynyamathanha we find *wa-nha(-)* ‘where?’, “consist[ing] of the interrogative *-wa* and the supporting element *-nha*” (Schebeck 1974:11-12), and that (iii) a tilde above a vowel sign would normally be expected to mark nasalization, it may be hypothesized that Teichelmann & Schürmann’s *ā* stands for something like *anh(a)*. Thus, *wā* ~ *wa* ‘where?’ is probably *wa-nh(a)* ~ *wa*. The possibility of a variant without *-nha* is explained by the fact that, as has been discussed above, *-nha* is likely to be a separate morpheme with some kind of specifying or focalizing function. The *-da* of *wāda* may be the same as the Adynyamathanha suffix *-tla* found e.g. on demonstratives instead of *-nha* and “convey[ing] the idea of lack of precision; it can also imply that the hearer is presumed to have some knowledge of the place referred to” (Schebeck 1974:11; see some examples above in the present section). The latter description fits well Teichelmann & Schürmann’s remark on the use of *wāda*. Finally, for the reasons already explained, Teichelmann & Schürmann’s demonstrative *inna* ‘this one’ is likely to be in reality something like *i-nha-nha*, its structure being the same as that of the identical Adynyamathanha demonstrative, already mentioned in the present section. Summing up, in all probability, Teichelmann & Schürmann’s interrogative *wādaina* ~ *wadana* ~ *wādanna* is phonologically something like *wa-nh(a)-da-i-nha* ~ *wa-da-nha* ~ *wa-nh(a)-da-nha-nha* respectively.

According to Hercus (1999:91), Wirangu has a general interrogative *ngana* ‘who?, what?’ (20), similar to the Thura-Yura languages Parnkalla and Kaurna (Section III.6.4.1.1.1).

Wirangu (Pama-Nyungan; WC)

- (20) a. *ngana-ngu gurnda-na?*
 IPW-ERG kill-PST
 ‘Who/what killed (him)?’ (Hercus 1999:91)
- b. *nyurni ngana-gu nhaa-na?*
 2SG IPW-DAT look-PST
 ‘Who/what did you look for?’ (Hercus 1999:91)
- c. *ngana-nga bala ngunyi-ri-rn?*
 IPW-LOC this.one laugh-REFL-PRS
 ‘Who/what is she laughing at?’ (Hercus 1999:91)
- d. *ngana nyurni yadu-ma-rn?*
 IPW 2SG good-make-PRS
 ‘What are you making?’ (Hercus 1999:77)

The interrogative *ngana* may “serve as indefinite pronoun when reduplicated”. Thus, *ngana-ngana* “means ‘whoever it might be’, ‘anybody’, but at the same time it can also mean ‘anything’” (Hercus 1999:91).

Furthermore, Hercus (1991:91-92) reports that besides “the basic interrogative pronoun” *ngana* ‘who?, what?’, Wirangu makes “a secondary distinction between ‘who’ and ‘what’” by means of *ngananha* ‘who?’ (21) and *nganhangga* ‘what?’ (22). Unlike *ngana*, the interrogatives *ngananha* and *nganhangga* are “found only in the absolutive form: [they] can NEVER be marked for case” (Hercus 1999:92).

Wirangu (Hercus 1999:92)

- (21) a. *ngananha nyurni?*
 IPW[ABS] 2SG
 ‘Who are you?’
- b. *ngananha urdli-na?*
 IPW[ABS] come-PST
 ‘Who came?’
- c. *nyurni ngananha nhaa-na?*
 2SG IPW[ABS] see-PST
 ‘Who did you see?’
- (22) *nganhangga bala-ardu?*
 IPW[ABS] this-FOC
 ‘What is this?’

However, the presumed restriction of *ngananha* to human use is contradicted by Hercus' (1999) own examples found in other sections of her Wirangu grammar. As to *nganhangga*, I have not found exact counterexamples, only an example with an almost identical form *nganhanga* meaning 'who?'.³¹

Let us start with an example where *ngananha* translated as 'what?', (23), which structurally is identical to (21c).

Wirangu (Hercus 1999:92)

- (23) *nyura ngananha nhaa-na?*
 2PL IPW[ABS] see-PST
 'What did you see?'

Thus, Wirangu *ngananha* appears to be used both as 'who?' and 'what?'. It is then identical to the Kurna 'who?, what?' interrogative *ngananha* (*nganna* in the original spelling) discussed in Section III.6.4.1.1.1. Hercus (1999:92) describes Wirangu *ngananha* as a "derivative of *ngana*". Apparently, it is not described as a case form of *ngana* because in Wirangu *-nha* is not as productive, as it is for instance in Adynyamathanha (cf. Section III.6.4.1.1.1).³¹

Hercus (1999:92) suggests that the Wirangu interrogative *nganhangga* 'what?' contains "another interrogative stem, *nganha*, which in [Adynyamathanha and Kuyani] has come to mean 'who'". I agree with Hercus that the two interrogatives are related, but I rather believe that Adynyamathanha and Kuyani *nganha* is just a contracted form of the already familiar interrogative *ngananha* 'who?, what?', which in these two languages has become specialized as 'who?' (just as *ngananha* has become specialized as 'what?' in Panyjima (WHc3; Dench 1991:164)). The contraction was probably preceded by an assimilation of *n* of the second syllable to the laminal articulation of *nh* of the last syllable, i.e. **ngananha* > **nganhanha* > *nganha*. In Wirangu *nganhangga* the last syllable is likely to represent the frozen locative/circumstantial/accompaniment suffix *-nga* (cf. Hercus 1999:53-55). Thus, in Wirangu the evolution has in all probability proceeded as follows: **ngananha-nga* > **nganhanha-nga* > *nganhangga*. The form *nganha-nga* in (24) suggests that alternatively, just like in Adynyamathanha, one of the two *nha* syllables of **nganhanha-nga* may be dropped.

Wirangu (Hercus 1999:54)

- (24) *dyirlbi warla nganha-nga?*
 old.man angry IPW-LOC
 'The old man is angry with whom?'

³¹ For instance, the segment *-nha* appears in Wirangu in some case forms of the first singular personal pronoun and is used to form certain demonstratives (cf. the discussion Hercus 1999:64-65, 72-84).

Note that *nganha-nga* ‘with whom?’ is almost identical to *nganhangga* ‘what?’.

Summing up, originally Wirangu appears to have had a single interrogative *ngana* ‘who?, what?’, which could be modified with the specific (or focus marker) *-nha*, similar to the Panyjima “specific referent marker” *-nha* already mentioned several times above. That is, the form *ngana-nha* presumably used to mean something like ‘who/what exactly?’. The present-day Wirangu language, as described by Hercus (1999), has preserved the interrogative *ngana* ‘who?, what? (can be used in any case)’, as well as some fossilized reflexes of *ngana-nha* ‘who?, what?’, such as *ngananha* ‘who?, what? (only ABS)’, *nganhangga* ‘what? (only ABS)’ and *nganha-* ‘who? (only LOC?)’ (perhaps, ‘what?’ as well, but examples are lacking). Although etymologically transparent, the fossilized forms should probably be considered unanalyzable from the synchronic point of view.

6.4.1.1.3 Mirning, Ngadjunmaya and Karlamay

The Pama-Nyungan languages Mirning/Mirniny (WE1), Ngadjunmaya/Ngatjunmaya/Kalaaku/Kalarko/Malba (WE2) and Karlamay (WE3) form the “Western Bight group” (WE) in Dixon’s (2002) classification. According to Dixon (2002:328), this is a group with a language/languages having a ‘who?, what?’ interrogative. Provided Dixon’s claim is correct, the respective interrogative is most likely to be formally similar to the Thura-Yura ‘who?, what?’ interrogatives of the *ngana(nha)* type discussed in Sections III.6.4.1.1.1–III.6.4.1.1.2. Thus, Dixon (2002:331–332) mentions the WE group among the groups with languages where “*nga:n-* [...] underlies ‘who?’”. Furthermore, according to the record on Mirning in Norman Tindale’s Catalogue of Australian Aboriginal Tribes,³² the language name of one of the major subdivisions of Mirning was “**Ngandatha**, based on the phrase ‘What is it?’” (emphasis added).

The only sources on these languages I was able to consult are two wordlists, O’Grady’s (1968) on Mirning and von Brandenstein (1980) on Ngadjunmaya. The available Mirning interrogative pronominals are reproduced in (25) and the Ngadjunmaya ones in (26). The abbreviations *qs*, *ps*, *p* in (25) must stand for the original source

Mirning (Pama-Nyungan; WE1; O’Grady 1968)

- | | | |
|------|---------------------|----------------------|
| (25) | <i>ngarntu</i> | ‘why (qs)’ |
| | <i>ngarntulu</i> | ‘who (ps)’ |
| | <i>ngarnturniny</i> | ‘what (p)/ who (qs)’ |

³² The catalogue is available at: <http://www.samuseum.sa.gov.au>.

Ngadjunmaya (Pama-Nyungan; WE2; von Brandenstein 1980)

- (26) a. *ngaandu* “interr[ogative].pron[oun]. who” [ABS?]
ngaandulu “interr.pron. + AG/INST [i.e., ERG/INS] by whom”
ngaan(d)ilu “interr.pron. + AG/INST [i.e., ERG/INS] by whom”
ngaandungarra “interr.pron. + PL who” [ABS?]
ngaanduwanja “interr.pron. + {*wanja*} whose” [GEN?]
- b. *wanja* “indef[inite]./interr.pron. (any) who, (any) where” [ABS?]
wanjina “indef./interr.pron. + ESS somebody, who” [ABS?]
wanjaa “indef./interr.pron. + LOC about anything else”³³
wanjardi “indef./interr.pron. + {*wardi*} about anything else”
wandanguu “indef./interr.pron. + {*nguu*} where from”
wanjipärrri “indef./interr.pron. + {*pärrri*} (any) where about”

It seems that *-tu* in (25) and *-du* in (26) is a frozen ERG/INS suffix. Thus, in Ngadjunmaya *-du* is the allomorph of the ERG/INS suffix used “on stems ending in cluster with {*n-*}”, as in *madarn-du* ‘by mother-in-law’ (von Brandenstein 1980). The instrumental function would explain the meaning ‘why?’ of *ngarntu* in (25). It is not clear whether Ngadjunmaya *ngaandu* may also be used as ‘why?’ or ‘what?’. I have found no other interrogative form meaning ‘what?’ in von Brandenstein (1980).

The polysemy ‘where?, who?’ with Ngadjunmaya *wanja* is not implausible. The original meaning here is clearly ‘where?’. The intermediate stage of the selective ‘which one?’ would be expected (cf. Section III.6.3.2). Unfortunately, no examples are provided to confirm the human use of *wanja*.

6.4.1.1.4 Pintupi

Pintupi (WDf), as described by Hansen & Hansen (1978), is a Pama-Nyungan idiom belonging to the so-called Western Desert (or Wati) language. It should be mentioned that linguistically, the divide between Pintupi and the neighbouring Western Desert language varieties is “anything but clear” (Hansen & Hansen 1978:17). The situation is further complicated by the existence of the so-called “hordelects, or vocabulary used mainly by particular family groups” and some other sociolinguistic factors and historical influences (cf. Hansen & Hansen 1978:17-25).

³³ The “ESS” in *wanji-na* “expresses identification either as copula = ‘is/are’ or as personification = PERS = ‘as the one’ or as GEO, identifying places (not translatable)” (von Brandenstein 1980). This *-n(a)/-ny(a)* is clearly cognate to the same “specificity”/focus marker *-nha/-nya* as the one in many other western and central Australian languages (cf. e.g. Sections III.6.4.1.1.1-III.6.4.1.1.2 and Sections III.6.4.1.1.4-III.6.4.1.1.5).

Hansen & Hansen (1978:196) mention two interrogative pronominal roots: (i) *ngana-*, as in *ngana-nya* ‘who-ABS?’, and (ii) *nyaa-*, as in *nyaa-lu* ‘what-ERG?’, *nyaa-tyanu-Ø* ‘what-ORIGIN-ABS?’, *nyaa-ku* ‘what-DAT?’. However, these glosses happen to be in contradiction with the glosses and translations of the sentential examples with *ngana-* and *nyaa-*, reproduced here in (27-29).³⁴ For ease of comparison, I have preserved the original glosses of *ngana-* and *nyaa-* used in these examples and I have highlighted those of them that are in contradiction with the glosses just mentioned.

Pintupi (Pama-Nyungan; WDf)³⁵

- (27) *ngana-nya yirrupulangka ngalya pityang-u?*
 who-ABS airplane.LOC toward come-PST.PUNCTILIAR
 ‘Who came in the plane?’ (Hansen & Hansen 1978:196)
- (28) *nyaa-lu wiya yaturn-u yilipi-ngka?*
who-ERG NEG chop-PST axe-LOC
 ‘Who didn’t chop him with the axe?’ (Hansen & Hansen 1978:196)
- (29) *ngana-ngka nyaa munta ngara-Ø-nytya-ngka?*
what-LOC **what** QUERY stand-PST.PUNCTILIAR-NMLZ-LOC
karapana-ngka nyukiniyi-ngka
 caravan-LOC PROP-LOC
 ‘[Yes, there should be a caravan with food there. Like the one which was at Lambra bore, a great big caravan.] What was it that was there? at what? a caravan? Oh. A New Guinea (type shelter). [We used to buy at a New Guinea shelter, at a very small opening in the New Guinea shelter.]’
 (Hansen & Hansen 1978:236, 238)³⁶

Only in (29) the use of *ngana* may be explained away as an example of ‘who?’-dominance in the case of a non-prototypical combination of values of the kind [thing + identification + proper name] (cf. Sections II.3.3.3 and II.3.3.6). In (29), the speaker is seeking the name of a named place (or a named object). Furthermore, note that strictly speaking this example is not a real question. It comes from a monologue and here the speaker apparently tries to remember the name of a shelter, *nyukiniyi* ‘New Guinea’. This implies that the roots *ngana* and *nyaa* in (29) are rather used as fillers, i.e. words like ‘whatchamacallit’ or

³⁴ No other examples of the use of the interrogatives *ngana-* and *nyaa-* have been found in the source.

³⁵ For ease of presentation, I have slightly modified the original glosses. Most importantly, I use LOC instead of Hansen & Hansen’s “accessory” and “general adjunct”.

³⁶ Hansen & Hansen (1978) gloss *munta* as QUERE, but this gloss is not in their list of glosses used (“table of symbols”). I putatively replaced QUERE with QUERY (it is not Hansen & Hansen’s list of glosses either).

‘whatsit’, so that the whole sentence should probably be better translated as ‘At whatsit... whatchamacallit... at a caravan... at a New Guinea’.

Unlike (29), example (28) cannot be explained by appealing to any kind of non-prototypical combination of values. Example (28) clearly suggests that *nyaa-* can also mean ‘who?’ next to ‘what?’. In the other Western Desert varieties I have data for, the interrogatives *nyaa/nhaa* seem to be used only as ‘what?’. The same is typical for most of the rest of the west and central Australian Pama-Nyungan languages, although there are also some Pama-Nyungan languages spoken to the north-east of Pintupi, where *nha-/nya-* is found to begin both the forms for ‘who?’ and for ‘what?’. In one case, Warumungu, we also find *nyayi* meaning both ‘who?’ and ‘what?’ (cf. Section III.6.4.1.1.5). I believe that in west and central Australian Pama-Nyungan languages this *nha-/nya-* interrogative goes back to the same “specific” (or focus) marker **-nha/*-nya* that frequently becomes frozen on their interrogative pronominals, as in Wirangu *ngananha* ‘who?, what?’ (WC; cf. Section III.6.4.1.1.2) and Panyjima *ngananha* ‘what?’ (WHc3; Dench 1991:164).³⁷ For some examples of the marker **-nha/*-nya*, cf. Sections III.6.3.4, III.6.4.1.1.1, III.6.4.1.1.5. The interrogative root *ngana* itself is found both as ‘who?’, as in Pintupi or Panyjima, and as ‘who?, what?’ in Yulparija, Walmajarri and Warlmanpa (cf. Section III.6.4.1.1.6). The development of *ngananha* to *nhaa* (and the like) will thus be an instance of clipping (with compensatory lengthening of the final vowel), a phenomenon which appears to be quite common with interrogatives in Australia (cf. Section III.6.3.4).

The predominant association of the *nha-/nya-* (and the like) interrogatives with the non-human meaning ‘what?’ in many west and central Australian Pama-Nyungan languages is in all probability due to the fact that the specificity/focus marker **-nha/*-nya* tends to be reanalyzed as an accusative or absolutive case marker, although usually restricted to proper names (both personal *and* non-personal), personal pronouns, demonstratives, interrogatives, and sometimes kin terms. Only very rarely does this suffix develop into “the marker of proper nouns, irrespective of their syntactic function in a clause” (Dixon 2002:155).³⁸

³⁷ The *nhaa/nyaa* ‘what?’ interrogatives of the Pama-Nyungan languages of the eastern part of Australia should in all probability be brought back to the root *minha/minya* ‘(edible) animal, something, what?’, which according to Dixon (2002:334) also “occurs almost exclusively in the eastern part of the continent”.

³⁸ Note that for Dixon (2002:155-156) the suffix *-nha/-nya* is “the pan-Australian form of the accusative suffix [that] typically occurs on pronouns”, with all functions, other than the marking of the accusative being later developments of this original function. I believe that the specificity/focus marking function is the primary one. The later association of *-nha/-nya* with the accusative/absolutive case marking function can be easily explained as a reanalysis by analogy. Given that OBJ and S (as well as predicate nominal) functions are the only functions

Warumungu (Simpson 2002:7.4)

- (32) a. *nyayi angi?*
 IPW you
 '[A:] Who are you? ([B:] I'm Nangali.)'
- b. *nyayi pangkal alinya kurarrapan?*
 IPW perhaps that run.here.PRS
 'Who could that be running here?'
- c. *nyayi-ngki aju kurarrapan?*
 IPW-ERG me poke.PST
 'Who poked me?'
- (33) a. *nyayi payinti?*
 IPW today
 '[A:] What's today? ([B:] It's Tuesday.)' or '[A:] What's it like today? ([B:] It's hot.)'
- b. *nyayi angi kupunta?*
 IPW you cook.PRS
 '[A:] What are you cooking? ([B:] Bush potato.)'
- (34) a. *nyangirr angi kupunta?*
 how you cook.PRS
 '[A:] How are you cooking it? ([B:] In the ashes.)'
- b. *nyangirr angi pinanta?*
 how you think.PRS
 'What do you think?'
- c. *nyangirr-ji-nyi Nampin tapinyila?*
 how-do-PST PROP morning
 '[A:] What did Nampin do this morning? ([B:] She went hunting.)'

Besides being an interrogative, the word *nyayi* 'can be used to mean 'someone' or 'something', especially before the word *pangkal* 'perhaps, maybe', or with the negative word *warra*. Examples (35a) and (35b) illustrate the latter two possibilities for the meaning 'something'.

Warumungu (Simpson 2002:7.4.2)

- (35) a. *pirntimunta arni nyayi pangkal kuyu*
 smell.PRS 1SG INDF perhaps meat
 'I can smell something, maybe meat.'
- b. *warra nyayi-nji jukunta, warraku, purnuku purtu*
 NEG INDF-ERG carry.PRS nothing car without
 'There's nothing to take (us) in, nothing, no car.'

Dixon (2002:328-329) suggests that the original meaning of *nyayi* was ‘what?’ “being extended also to have the sense ‘who’”. Apparently, Dixon’s hypothesis is based on the fact that, as he discusses elsewhere (2002:333-335), in the overwhelming majority of cases interrogatives beginning with *nha(:)/nya(:)-* mean ‘what?’ in Australian languages. In fact, *Australian* here may be virtually equated with *Pama-Nyungan*, because Dixon mentions only two such non-Pama-Nyungan languages, Bininj Gun-wok (Gunwinyguan; NBg1) and Jingulu (Mirndi group; NCb1). Furthermore, at least in Jingulu, the respective ‘what?’ interrogative looks very much like a Pama-Nyungan loan (cf. Section III.6.4.2.3).

According to Dixon (2002:333), there are just some four Pama-Nyungan languages where, like in Warumungu, we also find forms for ‘who?’ to begin with *nha-/nya-*. Kalkatungu (W1) ABS *nhani* and ERG *nhantu*, Yalarnnga (W2) ABS *nhanha* and ERG *nhantu*, Kaytety/Kayteye (WL2) *nhant*, and Djinba (Yc2) *nyalung*. Remarkably, of these four languages, one language, Kaytety/Kayteye (WL2), is spoken immediately to the south of Warumungu, and Kalkatungu (W1) and Yalarnnga (W2) are not so far away either. Only the last language, Djinba (Yc2), is spoken at a really substantial distance in a Pama-Nyungan enclave in the north-eastern Arnhem Land. Note also that at least in Kalkatungu, Yalarnnga and Djinba⁴⁰ the non-human interrogative ‘what?’ also begins with *nha-/nya-*, viz. *nhaka*, *nhangu* and *nhani* respectively (Dixon 2002:333). Dixon (2002:334) points out that the forms of ‘who?’ “in W and WL suggest a root **nhan*”. However, it may be safer to assume **nha-* ‘who?, what?’ instead, with various increments, which most likely go back to some frozen case markers or deictics.⁴¹ This **nha-* ‘who?, what?’ may have the same origin as *nyaa* ‘who?, what?’ in Pintupi (cf. Section III.6.4.1.1.4). That is, it may be related to the same “specific” marker **-nha/*-nya* that frequently becomes frozen on the interrogative pronominals of the west and central Australian Pama-Nyungan languages. In Warumungu, one also finds the “emphatic and deictic” suffix *-nya*, as in (36).

Warumungu (Capell 1953:306)

- (36) a. *ngala-nya angginyi wingara*
 DEM-EMPH 2SG.POSS totem
 ‘This is your totem (pointing and emphasizing).’

⁴⁰ I do not have information for Kayteye.

⁴¹ For instance, as far as *-yi* of *nyayi* is concerned, compare the “old Thura-Yura [Pama-Nyungan; WB] and Wirangu [Pama-Nyungan; WC] forms” of the “nuclear cases” of the first singular personal pronoun: *ngadhu* ‘1SG.A’, *ngayi* ‘1SG.S’, and *nganha* ‘1SG.ACC’ (Hercus 1999:73). The *-yi* of *nyayi* may also be compared to the Pitta-Pitta (WAa1) “distance” clitic *-yi*, which, according to Dixon (2002:305), is always added to the third person pronouns.

- b. *bani-nya* *gadila*
 like.this-EMPH make.IMP
 ‘Make it like this.’

In all probability, the original interrogative in Warumungu was based on the root *ngana* and was similar to Wirangu and Kaurna *ngananha* ‘who?, what?’ (cf. Sections III.6.4.1.1.1-III.6.4.1.1.2).

6.4.1.1.6 Yulparija, Walmajarri and Warlmanpa

Three languages from the northern fringes of the western and central parts of the Pama-Nyungan area, Yulparija/Yulparitja (WDb; Burridge 1996:32-33), Walmajarri/Walmatjarri (WJa1; Dixon 2002:328) and Warlmanpa (WJb3; Nash 1997; David Nash, p.c.), use a single interrogative pronominal *ngana* ‘who?, what?’. At least in Yulparija and Warlmanpa *ngana* can also be used attributively as ‘which [N]?, what [N]?’ and indefinitely as ‘someone, something’. I lack data on Walmajarri. Examples (37) and (38) illustrate the use of *ngana* ‘who?, what?’ in Yulparija and Warlmanpa respectively.

Yulparija (Pama-Nyungan; WDb; Burridge 1996:32-33)⁴²

- (37) a. *ngana nyarra?*
 IPW[ABS] DEM
 ‘Who/what is that?’
 b. *ngana-ku-nta-ra-n ngara-nyin?*
 IPW-DAT-Q-3SG.DAT-2SG.NOM be-PRS
 ‘What do you want?’
 c. *ngana-lu-nta ka-ngu jii-ngulyu mutuka?*
 IPW-ERG-Q carry-PST DEM-DEICTIC.ABS car[ABS]
 ‘Who drove that car?’

Warlmanpa (Pama-Nyungan; WJb3; David Nash, p.c.)

- (38) *ngana-rlu = ngku piya-rnu?*
 IPW-ERG=2SG.ACC bite-PST
 ‘Who/what bit you?’

Yulparija belongs to the so-called Western Desert (or Wati) language.⁴³ Probably with the exception of Pintupi (Section III.6.4.1.1.4) and maybe Lurija (cf. Section

⁴² The interrogative suffix *-nta* in the last two examples is one of the “three interrogative suffixes *-nta*, *-warapa*, *-rapa* which seem to optionally occur with *ngana* (and also *wanyja* [‘where?’])” (Burridge 1996:32).

⁴³ The Ethnologue classifies Yulparija/Yulparitja as a dialect of the Martu Wangka language of the Wati group.

III.6.4.1.2), most other Western Desert idioms do distinguish between ‘who?’ and ‘what?’. In these other idioms, the interrogative root *ngana* means ‘who?’.

Walmajarri and Warlmanpa belong to the Ngumpin-Yapa group of languages, Walmajarri being a Ngumpin language (Dixon’s WJa subgroup) and Warlmanpa a Yapa language (Dixon’s WJb subgroup). All other Ngumpin-Yapa languages⁴⁴ do distinguish between ‘who?’ and ‘what?’. All Ngumpin languages except Walmajarri have *ngana* ‘who?’ and *nyampa* (also written as *nyamba*) ‘what?’. Warlpiri, the only other language of the three Yapa languages I have data for, has *ngana* ‘who?’ and *nyiya* ‘what?’.

Adducing the aforementioned data on Ngumpin languages, Dixon (2002:328) suggests that in Walmajarri, as in most other languages with ‘who?, what?’ interrogatives, “the neutralisation [of ‘who?’ and ‘what?’] has been a recent, language-particular change”. However, judging from the facts discussed in the previous sections (Sections III.6.4.1.1.1-III.6.4.1.1.5), I believe that rather the situation must have been the other way around. That is, instead of the neutralization in the languages with a single ‘who?, what?’ interrogative there has been a specialization in the languages which now have a dedicated ‘who?’ and dedicated ‘what?’ interrogative. In this respect, recall first that *ngana*-based ‘who?, what?’ interrogatives happen to be rather widely distributed in the west and central Australian Pama-Nyungan languages. One finds them in the north and in the south of the latter area, while in the west we may find languages like Karajarri, where ‘what?’ is *ngana* (WJa2; Sands 1989:81-82), or Panyjima, where ‘what?’ is *ngananha* (WHc3; Dench 1991:164). Furthermore, “‘what’ appears to relate to *nga(:)n* in groups B-D, Ja and K [i.e., in the north-east of the continent], an area that lacks *nga(:)n* ‘who?’” (Dixon 2002:331). Second, recall that *nha-/nya-* ‘what?’ and ‘who?, what?’ interrogatives can be accounted for as going back to the *ngananha*-like derivatives of the same original ‘who?, what?’ interrogative *ngana* that have become specialized as ‘what?’ and clipped to *nhaa* and the like. In some languages, the latter clipped form has been subsequently incremented with other morphemes. For instance, the Ngumpin interrogative *nyampa* ‘what?’ in all probability contains the “hesitation”, “doubt” or “query” suffix *-mpa* common in the Pama-Nyungan languages of the area (cf. e.g. Eckert & Hudson 1988:25-26, 121 on Pitjantjatjara (WDm))

6.4.1.1.7 Mangala

Mangala/Mangarla (WIb) is a Pama-Nyungan language spoken immediately to the west of Yulparija and Walmajarri discussed in Section III.6.4.1.1.6. Dixon (2002) classifies Mangala as a member of the so-called Mangunj areal group (WI). The remaining members of this group are Nyangumarta (WIa1) and

⁴⁴ I only lack information on the Yapa language Ngardi (WJb2).

Karajarri (Wla2). According to Dixon (2002:328), the group WI contains a language that has a single form used for both ‘who?’ and ‘what?’. Given that Nyangumarta has *nganurtu* ‘who?’ and *ngani* ‘what?’ (Sharp 2004:256-257) and Karajarri has *ngardu* ‘who?’ and *ngana* ‘what?’ (Sands 1989:81-82),⁴⁵ we arrive to Mangala as the only language in this group that may contain such a ‘who?, what?’ interrogative. The only source mentioning Mangala interrogatives I have been able to consult is Nekes & Worms (2006). Unfortunately, the latter source does not mention any Mangala form for ‘who?’. Nekes & Worms (2006:142) only give Mangala *ngana* as ‘what?’. Nevertheless, given that (i) it follows from Dixon (2002) that Mangala has a ‘who?, what?’ interrogative, (ii) Mangala ‘what?’ is known to be *ngana*, and (iii) neighbouring Yulparija and Walmajarri have the same form *ngana* as both ‘what?’ and ‘who?’, it seems reasonable to conclude that in Mangala *ngana* can be used as both ‘what?’ and ‘who?’.

6.4.1.2 Some possible leftovers

Besides the Pama-Nyungan languages discussed in Section III.6.4.1.1, two more “languages” are mentioned by Schmidt (1919) that may prove to allow (or to have allowed) for a lack of differentiation between ‘who?’ and ‘what?’.⁴⁶ The first language is Lurija/Luritja (WDg), which is a Western Desert idiom. The second one is a group of languages or a dialect cluster that Schmidt calls “North Narrinyeri”, already extinct long time by now. As far as I can judge from the map of Australian languages in Schmidt (1919), his North Narrinyeri seems to correspond to Dixon’s languages U2 (Ngayawang), U3 (Yuyu or Ngarrket), U4 (Keramin, or Kureinji, etc.), and U5 (Yitha-Yitha, Dardi-Dardi). I reproduce Schmidt’s (1919) data in Table 3. For North Narrinyeri, Schmidt (1919) gives the genitive form *ḡannuno* in a footnote. The footnote number in the source is placed after the form *meyak* in the ‘who?’ cell. However, it is not clear whether this means that the genitive form of ‘who?’ is different from that of ‘what?’ or that the genitives of ‘who?’ and ‘what?’ are the same but simply based on a different root than the S and A forms.

Schmidt’s Lurija interrogative *yaal* is clearly a ‘where?’ interrogative in origin. In this respect, consider for instance (15) in Section III.6.3.4. A further extension of ‘where?’ to the meanings ‘who?’ and ‘what?’ through a selective ‘which one?’ would be semantically plausible and as such is attested elsewhere in Australia (cf. Sections III.6.3.2 and III.6.4.2.5). However, since no examples are provided, it is difficult to be sure that *yaal* really can be used in non-selective questions as both ‘who?’ and ‘what?’ (cf. a similar reservation concerning

⁴⁵ In Nekes & Worms (2006:142), Karajarri ‘who?’ is given as *ngadu*.

⁴⁶ Cf. also footnote 11 in Section III.6.2 on Schmidt’s “middle languages”.

Table 3. Interrogative pronominals of Lurija and “North Narrinyeri” (based on Schmidt 1919)

	‘who?’	‘what?’
Lurija	<i>yaal</i>	<i>yaal, ṇār, namba</i> ⁴⁷
“North Narrinyeri”	S-function: <i>meike, meyak</i> A-function: <i>meikenanna</i> GEN: <i>ṇannuṇo</i>	S-function: <i>meike, meyak</i> A-function: <i>meikenanna</i> GEN: ?

Kaurna *wadaina* in Section III.6.4.1.1.1 and Ngadjunmaya *wanja* in Section III.6.4.1.1.3).

6.4.2 Non-Pama-Nyungan languages

As expected, the non-Pama-Nyungan languages that appear to allow for one form to be used as both ‘who?’ and ‘what?’ show much more variation in the forms of the respective interrogatives than the Pama-Nyungan languages. Their discussion will be organized as follows. I will begin in Section III.6.4.2.1 by discussing Anindilyakwa, where the apparent lack of differentiation between ‘who?’ and ‘what?’ seems to be due to the form of gender-number markers applied to a single interrogative pronominal root. In Section III.6.4.2.2, I will examine the languages where the apparent lack of differentiation between ‘who?’ and ‘what?’ is due to the peculiarities of the semantics of the gender-number markers applied to an interrogative pronominal root, which as such is indifferent to the human vs. non-human distinction. Three such languages, Burarra, Worrorra and Wambaya, will be considered in detail. Next, I will discuss Wambaya’s neighbour and presumed distant relative Jingulu (a member of Mirndi language group; Section III.6.4.2.3). The sources disagree on whether it has a general ‘who?, what?’ interrogative, but if Jingulu does have such an interrogative it is most likely to be due to an extension of a selective interrogative (itself, in all probability, originally a locative ‘where?’) to non-selective contexts. After Jingulu, three non-Pama-Nyungan language families will be discussed where the lack of differentiation between ‘who?’ and ‘what?’ should apparently be reconstructed for the respective proto languages. The three families are: Nyulnyulan (Section III.6.4.2.4), Tangkic (Section III.6.4.2.5) and Garrwan (Section III.6.4.2.6).

⁴⁷ In all probability, Schmidt’s form *namba* is in reality *nyamba* (or *nyampa*, depending on the spelling), since in the Western Desert idioms (WD) the interrogative ‘what?’ usually has the form *nyaa*. Lurija *nyamba* ‘what?’ can be further compared to *nyamba* ‘what?’ in most Ngumpin Pama-Nyungan languages (Dixon’s group WJa), such as Mudbura (WJa4) or Jaru/Djaru (WJa2).

Finally, in Section III.6.4.2.7 I will discuss the Gunwinyguan language Rembarrnga, for which it has also been claimed that it is possible to use one and the same form as both ‘who?’ and ‘what?’, but where an interpretation in terms of a non-prototypical combination of values may prove to be more appropriate.

6.4.2.1 Form of the gender-number markers: Anindilyakwa

Anindilyakwa/Aninhdhilyagwa/Yingguru (NBd3) is a non-Pama-Nyungan language. It is currently viewed as an isolate, but may be distantly related to Gunwinyguan languages (cf. Evans 2003a:13). Capell (1942:376-379) describes it as having nine genders, some of which are exclusively human. Anindilyakwa also obligatory marks gender on the (non-selective) interrogative pronominal stem *-miebina*. Since both the prefix of the human plural gender G3 and the prefix of the non-human gender G9, which is the gender of “the larger animals” (including birds), have the same form *wura-*, the interrogative pronominal form *wura-miebina* may refer both to (plural) humans as ‘who?’ and (singular or plural) non-humans as ‘what?’.⁴⁸

An even more extensive syncretism of the human/non-human gender-number markers is found in the neighbouring Gunwinyguan language Nunggubuyu (NBd2; cf. Heath 1984:160), which is only distantly (if at all) related to Anindilyakwa. However, apparently unlike in Anindilyakwa, in Nunggubuyu “in the complete word forms these human/nonhuman syncretisms are usually differentiated by the co-occurrence of specifically human number-markers” (Heath 1984:160).

6.4.2.2 Gender-number semantics

In at least three non-Pama-Nyungan languages, Burarra, Worrorra and Wambaya, it appears to be possible to use the same form as both ‘who?’ and ‘what?’ and this possibility seems to be due to the peculiarities of the semantics of the gender-number markers applied to an interrogative pronominal root, which as such is indifferent to the human vs. non-human distinction (cf. Section III.6.3.3). Other languages that might belong here as well are Ngandi (non-Pama-Nyungan, Gunwinyguan; NBd1), Gurr-goni/Guragone (non-Pama-Nyungan, Maningrida; NBf2) and Jawoyn (non-Pama-Nyungan, Gunwinyguan; NBh1).

⁴⁸ Note that normally, I would not consider the use of the same interrogative in questions about humans and animals as lack of differentiation between ‘who?’ and ‘what?’ (Sections I.4.2.3.2, II.4.1.3). However, I do treat the Anindilyakwa interrogative *wura-miebina* as a ‘who?, what?’ interrogative because unlike e.g. in Russian, in Anindilyakwa the use of one and the same interrogative is not due to semantic considerations but due to (accidental?) formal identity of one of the human forms of the interrogative pronominal and one of its non-human forms.

However, since I have not been able to check the sources on whether in these languages the masculine and feminine genders also contain inanimates in addition to humans, they will not be discussed.

6.4.2.2.1 Burarra and Worrorra

Burarra (NBf1) is a non-Pama-Nyungan Maningrida language, other languages of this family being Gurr-goni/Guragone (NBf2), Nakkara (NBf3), and Ndjébbana/Djeebbana/Gunavidji (NBf4). Worrorra is a non-Pama-Nyungan Worrorran language (NG1, including five other “dialects”, Dixon 2002:xli), other languages of this family being Ngarinyin/Ungarinjin (NG2, including seven other “dialects”, Dixon 2002:xli) and Wunambal (NG3, including six other “dialects”, Dixon 2002:xli).

Both Burarra and Worrorra have two human genders, masculine and feminine, next to two exclusively non-human genders. The meanings ‘who?’ and ‘what?’ are expressed in Burarra with one and in Worrorra with two interrogative pronominal stems marked by the appropriate gender marker. The interrogative pronominal stem in Burarra is *-(yi)nga* (Glasgow 1984:42). Worrorra has *angku-* for the masculine and the feminine and *angu-* for the two non-human genders (Clendon 1994 via Dixon 2002:476; Love 2000:16-17), as illustrated in (39).

Worrorra (non-Pama-Nyungan, Worrorran; NG1; Love 2000:16)

(39) G1 (masculine)	<i>angku-yi</i>	‘who?’
G2 (feminine)	<i>angku-nya</i>	‘who?’
G3 (non-human: places, etc.)	<i>angu-ja-ma</i>	‘what?’
G4 (non-human: residue)	<i>angu-ja</i>	‘what?’

However, as was pointed out in Section III.6.3.3, nor in Burarra nor in Worrorra are the two human genders restricted exclusively to nouns with human or even animate referents. What is more, the inanimate nouns appear to be assigned to the masculine or the feminine gender following some clear semantic principles. For instance, in Worrorra, “all tools used by the white man are described as masculine”, similarly many objects produced or strongly related to men appear to be masculine, while shiny things, such as glass, seem to end up in the feminine gender because they “shine like the sun” (Love 2000:16, 22). Somewhat similarly, in Burarra the masculine gender includes, among other things, “metal objects” (Glasgow 1984:7). It is not implausible then that, for instance, if a Worrorra speaker sees an unknown “white man tool” or a Burarra speaker sees an unknown “metal object”, s/he would use the masculine gender form of the respective interrogative pronominal stem, i.e. *angku-yi* (Love 2000:16)/ *angku-yu* (Clendon 1994 via Dixon 2002:476) and *an(a)-nga* (Glasgow 1984:42) respectively. But this masculine form is also the default human interrogative

‘who?’. Unfortunately, the data available do not provide information on this point.

6.4.2.2.2 Wambaya

Wambaya (NCb3) is a non-Pama-Nyungan language belonging to the Mirndi language group, which further includes Ngarnka (NCb2), Jingulu (NCb1), Jaminjung (with Jaminjung and Ngaliwurru dialects; NCa1), and Nungali (NCa2) (cf. Section III.6.4.2.3 on Jingulu). Wambaya is most closely related to Ngarnka. The exact nature of their relation to the other Mirndi languages is not clear for the moment (cf. Harvey, Green & Nordlinger 2006:292). There are three Wambaya dialects: Gudanji, Bibinka and Wambaya proper. This section is dedicated to Wambaya proper, as described by Nordlinger (1993).

Wambaya has four genders, which can be conveniently labelled as masculine for G1, feminine for G2, vegetable for G3, and neuter (or residue) for G4 (cf. Nordlinger 1993:4.2). All animates (including humans) are in G1 and G2, of which G1 appears to be the most functionally unmarked gender. The majority of inanimates are assigned to G3 and G4. Wambaya uses a single interrogative pronominal root, which is obligatory marked for gender and case. The singular forms, as “found in the corpus”, are reproduced in Table 4.

Table 4. Singular forms of the Wambaya interrogative pronominal (based on Nordlinger 1993:4.7.1)

	ABS	LOC/INS/ERG	DAT
G1 (M)	<i>gayini</i>	<i>gayini-(ni)-ni</i>	<i>gayini-ni-nka</i>
G2 (F)	<i>gayini-rna</i>	<i>gayini-nga-ni</i>	<i>gayini-nga-nka</i>
G3 (VEG)	—	—	—
G4 (N)	<i>gayina</i>	<i>gayina-ni</i>	—

Mostly depending on the gender, the interrogative means ‘who?’ or ‘what?’. When it is used attributively, as ‘which [N]?, what [N]?’, it agrees in gender with the controller. The masculine interrogative pronominal *gayini* is the default form in questions about humans and thus the closest equivalent of ‘who?’. However, the masculine form is also used in “no conjecture”-contexts, viz. “when the referent is unknown, such that it is not possible to determine the gender” (Nordlinger 1993:4.7.1; cf. example (13) above). As has been mentioned in Section III.6.3.3, this does not count as a case of lack of differentiation between ‘who?’ and ‘what?’. Finally, as discussed in Section III.6.3.3, Wambaya appears to use *gayini* in questions about animals (cf. also Section II.4).

As already mentioned, in Wambaya the masculine and the feminine genders also contain some inanimates, most of which “refer to natural events or celestial bodies”, such as *jinkiji* ‘star (M)’, but one also finds nouns like *ginguli* ‘hook (M)’ (Nordlinger 1993:4.2.1). Therefore, it is reasonable to expect that the masculine or correspondingly the feminine interrogative pronominal will be used in questions about such inanimates, which would count as a lack of differentiation between ‘who?’ and ‘what?’. In this respect, compare also Burarra and Worrorra discussed in Section III.6.4.2.2.1.

Furthermore, since Wambaya does not mark gender with the dual number suffix, in the dual there is only one form of the interrogative pronominal for all genders, indifferent to the human vs. non-human distinction. The dual interrogative pronominal is *gayini-yulu* in the absolutive and *gayini-yuliji* in the ergative/locative/instrumental (Nordlinger 1993:4.7.1).

By way of conclusion, it may be worth mentioning that in Wambaya the forms used as interrogatives can also be used as indefinites, although in the latter function they are “generally marked with the inferential clitic =*miji*” (Nordlinger 1993:4.7.6). Presumably, the dual form *gayini-yulu*, which is indifferent to gender and consequently to the opposition human vs. non-human, may also be used in the indefinite function about both two people and two things.⁴⁹ However, Nordlinger (1993) does not explicitly mention this possibility.

6.4.2.3 Jingulu

Jingulu/Djingulu/Jingili/Djingili (NCb1) is a non-Pama-Nyungan language usually viewed as a member of the Mirndi language group, which further includes Wambaya (with Wambaya, Gudanji and Bibinka dialects; NCb3), Ngarnka (NCb2), Jaminjung (with Jaminjung and Ngaliwurru dialects; NCa1), and Nungali (NCa2). However, it is not clear for the moment whether Proto Mirndi has ever existed. In other words, Jingulu may be related to the rest of the Mirndi group through some “higher-level proto-language”. (cf. Harvey, Green & Nordlinger 2006:292). There is a disagreement between the sources on whether in Jingulu a single form can be used both as ‘who?’ and ‘what?’. Thus, the existence of such a form is argued by Pensalfini (1997, 2003), but not by Chadwick (1975). Let us discuss the two points of view in more detail, starting by Pensalfini (1997, 2003). At the end of this section, some etymological notes will be made.

⁴⁹ Possible contexts would be, e.g., *I’ve just seen some two persons on the top of that hill* for the human meaning and *When I was sitting under that tree yesterday, some two things fell on my head from the tree* for the non-human meaning.

According to Pensalfini (1997, 2003), Jingulu has one dedicated ‘what?’ interrogative pronominal, *nyamba* (40, 42a), and one ‘who?, what?, which?’ interrogative pronominal, *(w)aja/(w)aji* (41, 42).

Jingulu (non-Pama-Nyungan, Mirndi group; NCb1)

- (40) *nyamba jimi-niki-(r)ni? jimi-niki-(r)ni darrangku*
 what DEM.N-PROXIMAL-FOC DEM.N-PROXIMAL-FOC tree
ngaba-nya-ju
 have-2SG-PRS
 ‘What’s that? That’s a tree you have there’ (Pensalfini 1997:237, 2003:142)
- (41) *aji-rni-mbili ya-miki jama-(r)ni-ma?*
 IPW-FOC-LOC 3SG-PST.VENTIVE DEM.M-FOC-EMPH
 ‘Who/what did he bring?’ (Pensalfini 1997:239, 2003:144)⁵⁰
- (42) a. *aja-rni nyamba ngaba-miki jama-niki-rni?*
 IPW-ERG what have-PST.VENTIVE DEM.M-PROXIMAL-ERG
 ‘Who among these people brought what?’ (Pensalfini 1997:200, 2003:101)
- b. *waju aji-rni bundurru-nu ngini-niki ngaba-miki?*
 IPW IPW-ERG food-PST DEM.N-PROXIMAL have-PST.VENTIVE
 ‘Who brought what of this food?’ (Pensalfini 1997:200, 2003:101)

The final *-a* of *(w)aja* may optionally become *-i* before certain suffixes. It may also become *-u* before *w*, as in *waju-wa* ‘where?’, presumably derived through the addition of the itive future marker *-wa(ra)* (Pensalfini 2003:144). Apparently, in (42b) the form *waju* is also due to an assimilation to the lost initial *w-* of *(w)aji-*. Furthermore, word-initial glides, *w-* and *y-*, are often dropped in Jingulu (Pensalfini 1997:63-65), which explains the variation *waja ~ aja* (and *waji ~ aji*).

Both *nyamba* and *(w)aja/(w)aji* can take on the indefinitivizer *-nayi* to form the indefinite pronominals *nyamba-nayi* ‘something’ and *(w)aja-nayi/(w)aji-nayi* ‘something, somebody’ (Pensalfini 1997:204, 297, 542, 547). The ergative suffix may be inserted between *(w)aja-/(w)aji-* and *-nayi*. Normally, the resulting forms *(w)aji-ni-nayi/(w)aji-li-nayi/(w)aji-yi-nayi* appear to mean only ‘someone’ (Pensalfini 1997:204, 547).

Chadwick (1975:39-40 via Mark Harvey, p.c.) analyzes the two interrogative pronominals of Jingulu somewhat differently, viz. he defines *waja* as ‘who?’,

⁵⁰ Pensalfini (1997) glosses *-rni-* in *aji-rni-mbili* as FOC but in (2003:145) he states: “The */-rni/* that appears on *(w)aja* in [this] example [...] is neither Ergative nor Focus marking, though it is homophonous with both. Its appearance in this word does not appear to have a morphological or semantic rationale”.

which?’ and *nyamba* as ‘what?’. Mark Harvey (p.c.) argues that “the examples Pensalfini provides (2003:142-145) fit with this distinction”, viz. “it appears that all of the *waja* questions translated with ‘what?’ could equally be translated with ‘which (one)’”. He further points out that Chadwick worked with fluent speakers whereas Pensalfini worked with semi-speakers, which some might interpret in favour of Chadwick’s analysis. However, in principle, the two analyses are not irreconcilable. If the selective meaning ‘which one?’ of *waja* were taken as the diachronically primary, one could easily imagine how it may have been expanded to non-selective contexts, but that this expansion was not fully simultaneous for the questions about humans and non-humans. Both semantically and typologically, the link between ‘which one?’ and ‘who?’ appears to be stronger than that between ‘which one?’ and ‘what?’, which would make it plausible that *waja* ‘which one?’ has been first expanded to non-selective contexts involving humans, i.e. as ‘who?’, and only later to non-selective contexts involving non-humans, i.e. as ‘what?’.

Let us now briefly discuss some further etymological issues related to the two interrogatives, starting with *nyamba*. Pensalfini (1997:236, 2003:142) suggests that *nyamba* ‘what?’ is “probably related to the /nyam-/ demonstratives”, (43).

Jingulu (Pensalfini 2003:133)

- (43) M *nyama*
 F *nyamarni*
 N *nyambala*
 VEG *nyambala*

Interestingly, I found example (44), where the meaning ‘what?’ is expressed with *nyambala*, which is also the form of the “/nyam-/ demonstrative” in the neuter and vegetable genders.

Jingulu (Pensalfini 1997:200, 2003:100)

- (44) *nyambala ngaja-nya-ju?*
 what see-2SG-PRS
 ‘What do you see?’

This form of ‘what?’ is not otherwise explicitly mentioned by Pensalfini (1997, 2003). It is worth mentioning that in the dictionary included in Pensalfini (1997) one also finds *nyambala* “(Demonstrative (Indefinite) n [neuter], v [vegetable]) ‘a, some’” (1997:542), as in (45).

Jingulu

- (45) *ibilkini-mbili yurrayu-ju wawa nyambala karalu*
 rain-LOC play-PRS child.M DEM.N earth
 ‘During the rains the [sic!, some?] children play in mud’ (Pensalfini 1997:542)

Another possibility is that the Jingulu interrogative *nyamba* ‘what?’ is a Pama-Nyungan loan. Compare, for instance, *nyamba* ‘what?’ in the Pama-Nyungan languages of the Ngumpin group, e.g. Mudbura (WJa4; David Nash, p.c.), adjacent to Jingulu, and Jaru/Djaru (WJa2; Tsunoda 1995:63), spoken somewhat further to the west. The ‘what?’ interrogatives of the form *nyaa* and the “hesitation”, “doubt” or “query” suffix *-mpa* are also common in Western Desert idioms (WD) (cf. e.g. Eckert & Hudson 1988:25-26, 121 on Pitjantjatjara (WDm)). The form *nyambala* ‘what?’ in (44) might then be accounted for as the result of the reinterpretation of the borrowed *nyamba* ‘what?’ as being related to the neuter/vegetable gender form of the “/nyam-/ demonstrative”.

Similarly to *nyamba* ‘what?’, the Jingulu interrogative *(w)aja/(w)aji* ‘who?, what?, which (one)?’ may have also been borrowed from Pama-Nyungan. In particular, it may be related to Pama-Nyungan ‘where?’ interrogatives such as Bularnu *wadha* (WMb2; Breen 2003:441), with a normal semantic change from ‘where?’ to ‘which (one)? (person or thing)’ and ultimately to ‘who?, what?’.

6.4.2.4 Nyulnyulan languages

Apart from Ngumbarl and to a certain extent Bardi, all Nyulnyulan languages⁵¹ have a single interrogative word ‘who?, what?, which [N]?, what [N]?’ which “covers not just persons and things but also events” (McGregor 2004:128). For Ngumbarl there is simply “virtually no material” available (cf. Stokes & McGregor 2003:33). In Bardi, the situation is a bit more complicated than elsewhere in Nyulnyulan and will be presented further below. Before proceeding to the Bardi data, some relevant patterns of the Nyulnyulan interrogative pronominal systems will be illustrated, mainly on the example of Yawuru, Nyulnyul and Warrwa. The section will be concluded by some etymological notes.

Yawuru, as described by Hosokawa (1991:495-499), uses one general interrogative pronominal *yangki* ‘who?, what?’ (46), which can also be used attributively, as in (47).

⁵¹ The remaining Nyulnyulan languages are Jabirrjabirr, Jawi, Jukun (this maybe a dialect of Yawuru), Nimanburru, Nyikina (with Big an Small Nyikina dialects), Nyulnyul (with Coastal and Inland Nyulnyul), Warrwa, Yawuru (with Julbayi and Marangan dialects) (Stokes & McGregor 2003:29-31; cf. also Section III.6.2 on the number of Nyulnyulan “languages”).

Yawuru (non-Pama-Nyungan, Nyulnyulan; NE1)

- (46) a. *yangki i-m-bula-rn miliya?*
 IPW[ABS] 3MIN-EP-come-IPFV now
 ‘Who has arrived?’ or ‘What has been delivered?’ (Hosokawa 1991:495)
- b. *yangki-barri i-rndira-nda ginyangka?*
 IPW-INS 3MIN-go-PFV 3MIN[ABS]
 ‘With whom did he go?’ or ‘What did he bring?’ (Hosokawa 1991:497)
- c. *yangki-yi nyamba marrkirdi?*
 IPW-DAT this hat[ABS]
 ‘Whose hat is this?’ or ‘What is this hat for?’ (Hosokawa 1991:498)
- (47) *yangki maya junggarra (i-nga-rn)?*
 IPW house[ABS] 2AUG.DU.GEN 3MIN-be-IPFV
 ‘Which house is yours?’ (Hosokawa 1991:495)

In some cases, *yangki* may be much more likely to have the human reading ‘who?’, e.g. the ergative *yangki-ni* (or contracted *yagani*) (48a), in some the non-human reading, e.g. the causal *yangki-nyurdany* (48b), while in some the human or the non-human reading depends more on the larger context, as in (46). The human meaning ‘who?’ can be made explicit by means of ‘the phrase *yangki ngarrungu* (lit. ‘which people’) or its contracted form *yanggarru*’ (Hosokawa 1991:496).

Yawuru

- (48) a. *yangki-ni rdii i-na-ra-nda nyamba?*
 IPW-ERG break 3MIN-TRANSITIVE-AUX-PFV this[ABS]
 ‘Who broke this?’ (Hosokawa 1991:496)
- b. *yangki-nyurdany ku-rry-jali-nda ngurdirn?*
 IPW-CAUSAL 2AUG-AUG-return-PFV alone
 ‘Why did you (PL) come back by yourselves?’ (Hosokawa 1991:498)

The interrogative *yangki* “does not refer to a place”, instead “[t]he locative interrogative *jana* ‘where?’ is used”, as in (49a). “When *yangki* does inflect for a local case, it is interpreted as indicating a continuous action in which the subject is engaged [as in (49b)] rather than location” (Hosokawa 1991:496).

Yawuru (Hosokawa 1991:497)

- (49) a. *jana mi-nga-rn?*
 where 2MIN-be-IPFV
 ‘Where are you?’ (Hosokawa 1991:496)

b. *yangki-gun mi-nga-rn?*

IPW-LOC 2MIN-be-IPFV

‘What are you doing?’ (lit.: ‘What are you in?’)

Interestingly, according to McGregor (2004:128), “in the Nyulnyul of the last full speaker, [the cognate general interrogative] *angki* was also almost always used for places as well, instead of [the dedicated ‘where?’ interrogative] *arr*”. However, the latter statement is in contradiction with McGregor’s own description of Nyulnyul (1996:18): “The determiner *angk* [i.e., the interrogative ‘who?, what?'] is also frequently used [in Nyulnyul] in requests for spatial location, direction, etc.; when used in this way it is *almost always followed by the nominal bur* ‘place, country’, which appears to be *almost compounded with it*” (italics added). Similarly to Yawuru, in Nyulnyul the human ‘who?’ can be made explicit by means of a phrasal constructions combining the general interrogative pronominal with a generic noun *wamb* ‘person, man’, as in *angka wamb* ‘what person?, who?’, although this construction is used only occasionally (cf. McGregor 1996:18). It is reasonable to expect similar constructions to be possible in other Nyulnyulan languages as well. However, nothing can be said on how common their use might be in a given language.

In some Nyulnyulan languages, the same form can be used as both the general interrogative ‘who?, what?’ and as an indefinite, both human ‘someone’ and non-human ‘something’. For instance, this is possible for Warrwa (*y*)*angki* (McGregor 1994:18). However, this is not allowed for the cognate Nyulnyul form *angk*: “Instead, indefinite meaning is expressed by means of general nouns such as *wamb* ‘man, person’ and *urany* ‘woman’ [, although] as in English [*angk*] can be used in the manner of relative pronominal[...] or complementiser[...]” (McGregor 1996:19). In Yawuru, *yangki* can be used as an indefinite only when marked by “the uncertainty-marking clitic *-marda*” (Hosokawa 1991:505-506, 581-582), as in (50).

Yawuru (Hosokawa 1991:506)

(50) *yangki-marda i-na-maku-makura-rn jimbin*

INDF-perhaps 3MIN-TRANSITIVE-RED-make-IPFV inside

‘I don’t know what he is doing inside (i.e. cooped up in his cut).’

Unfortunately, Hosokawa (1991) does not provide other examples of the indefinite *yangki-marda*. Therefore, one can only presume that it can be used about persons as well. This assumption is corroborated by the fact that Hosokawa (1991) does not explicitly mention any restrictions on the human use of this indefinite and by the fact that in Warrwa, which is rather closely related to Yawuru, the cognate form (*y*)*angki* can be used both as ‘something’ and ‘someone’. I do not have information on the other Nyulnyulan languages.

Let us now consider the interrogative pronominals of Bardi. According to Claire Bower, ⁵² Bardi is the only Nyulnyulan language that has “separate words for ‘who?’ *anggaba* and ‘what?’ *anggi*, but *anggaba* is a transparent cleft construction and must be really recent”. However, “even in Bardi the distinction is neutralised if the referent is clear from context, e.g. you can say [(51a)] or [(51b)] with no large difference in meaning”.

Bardi (non-Pama-Nyungan, Nyulnyulan; NE2; Claire Bower)

- (51) a. *anggaba aarli inamboona*
 b. *anggi aarli inamboona*
 IPW fish speared
 ‘Who speared the fish?’

Bower further believes that in all probability (although she “ha[s]n’t tried it”), *anggaba* can also be used as ‘what?’ when the context is clear enough, e.g. in the context like *I don’t believe you said you ate turtle brains. WHAT did you eat?*. However, the latter supposition would not seem as readily plausible to me as it appears to be to Bower. The reason is that the Bardi interrogative *anggaba* looks very much like a contraction of [**yangki* ‘IPW’ + **wamba* ‘man, person’], i.e. ‘what/which person?’, similar to Nyulnyul *angka wamb* mentioned above. ⁵³ Compare also the structurally similar Yawuru interrogative *yanggaru* < *yangki ngarrungu* ‘who?’ (lit. ‘which people?’). Summing up, it appears that Bardi has recently developed a dedicated human interrogative pronominal *anggaba* ‘who?’ from the older general interrogative pronominal *anggi* ‘who?, what?’, with the result that the latter now tends to be restricted to non-human uses as ‘what?’, although it can still be used as ‘who?’ as well.

By way of conclusion, let us consider some etymological issues related to the general ‘who?, what?’ interrogative pronominal of the Nyulnyulan languages, starting with its form. Thus, according to Stokes & McGregor (2003:67), the Proto Nyulnyulan ‘who?, what?’ interrogative should be reconstructed as **yangki*. This reconstruction may be supported by such forms as Yawuru and Nyikina *yangki* (McGregor 2004:128; Nekes & Worms 2006:138), Warrwa

⁵² Claire Bower (2003) at <http://listserv.linguistlist.org/cgi-bin/wa?A2=ind0303A&L=LINGTYP&P=R521>.

⁵³ The “loss of initial *w*” is a fairly common process in Bardi and Jawi, as in **wamba* ‘man’ > Bardi and Jawi *amba* (Stokes & McGregor 2003:56). The “loss of *m*” is also attested in Bardi and Jawi, although as a “minor and irregular process”, **karrambal* ‘bird’ > Bardi *garraabal*, Jawi *karraabal* (Stokes & McGregor 2003:59).

(y)*angki* (McGregor 1994:18), Jabirrjabirr, Nimanburru, Bardi and Nyulnyul *angk* (Nekes & Worms 2006:138).^{54, 55}

Since the lack of differentiation between ‘who?’ and ‘what?’ in Nyulnyulan appears to go back to the Proto Nyulnyulan period or even earlier, any hypothesis on the origins of this lack of differentiation is very likely to involve some other non-Pama-Nyungan languages or the neighbouring Pama-Nyungan languages. Naturally, the first non-Pama-Nyungan candidates to try would be the neighbouring Worrorran and Bunaban languages. These would be useful to look at even if ultimately unrelated to Nyulnyulan, because of a higher possibility of borrowing due to the long-term contact. Further away, Stokes & McGregor (2003:61) mention some “intriguing [...] morphological correlations” with the languages of the Mirndi group (Dixon’s NC group; cf. Sections III.6.4.2.2.2, III.6.4.2.3).

I have not found immediate parallels to Nyulnyulan among Bunaban languages or the neighbouring Pama-Nyungan languages, but the following Worrorran data may be of some interest. Worrorra has two human genders, masculine and feminine, next to two exclusively non-human genders. The meanings ‘who?’ and ‘what?’ are expressed with two interrogative pronominal stems marked by the appropriate gender marker. The stems are *angku-*, for the masculine and the feminine genders, and *angu-*, for the two non-human genders (Clendon 1994 via Dixon 2002:476; Love 2000:16-17). The respective gender forms of the two interrogative pronominal stems are summarized in (52).

Worrorra (non-Pama-Nyungan, Worrorran; NG1; Love 2000:16)

(52)	G1 (masculine)	<i>angku-yi</i>	‘who?’
	G2 (feminine)	<i>angku-nya</i>	‘who?’
	G3 (non-human: places, etc.)	<i>angu-ja-ma</i>	‘what?’
	G4 (non-human: residue)	<i>angu-ja</i>	‘what?’

Compare also the interrogative pronominals *nyanggi* ‘who?’ and *anja* ‘what?’ of Ngarinyin, another Worrorran language (NG2; Coate & Oates 1970:33). These Ngarinyin interrogatives look rather similar to the Worrorra masculine G1 and non-human residue G4 forms *angku-yi* and *angu-ja* respectively. The only major difference is the initial *ny-* in the Ngarinyin form for ‘who?’.

⁵⁴ Nekes & Worms (2006:138) spells *yangki* as *yangi* and *angk* as *ang*. This difference is just a matter of spelling conventions. Yet another spelling possibility for *angk/ang* would be *angg*.

⁵⁵ In Nyulnyul, next to the usual form *angk* ‘who?, what?’ (McGregor 1996:18), with a regular loss of the root final vowel (cf. Stokes & McGregor 2003:54), one also finds *angka wamb* ‘what person?, who?’, *angka-bur-ung* ‘where to? (IPW-place-ALL)’ or *angka yarrad* ‘[they are listening to] what we [are saying]’ (McGregor 1996:18-19). This *-a* is probably an epenthetic vowel inserted to avoid the consonant cluster.

From a purely formal point of view, the Proto Nyulnyulan interrogative **yangki* ‘who?, what?’ may very well be cognate to the Worrorra and Ngarinyin masculine interrogative pronominals *angku-yi* and *nyanggi*. Consider in this respect the variation in the initial consonant of the Yawuru proximal demonstrative: *nyamba* ~ *jamba* ~ *yamba* (Hosokawa 1991:472). Furthermore, note that although **ny* makes part of the Proto Nyulnyulan consonant inventory, no lexeme with an initial **ny-* is reconstructed for Proto Nyulnyulan, except for the verbal root **-NYA* ‘get, catch, pick up’ (cf. Stokes & McGregor 2003:38, 61-67).

Semantically, the link between the masculine interrogative and the general interrogative ‘who?, what?’ would not be unnatural for the non-Pama-Nyungan languages either. This is due to the fact that the masculine agreement pattern is often the functionally default agreement pattern in these languages, i.e. on certain targets and in certain contexts it can be used instead of any other agreement pattern. Consider, in this respect, Section III.6.3.3, as well as Section III.6.4.2.2.2 on Wambaya and Pensalfini (1997:253-263) on Jingulu. It may be worth pointing out as well that, as discussed in Section III.6.4.2.2.1, in certain contexts the Worrorra masculine interrogative *angku-yi* ‘who?’ is also likely to be used in questions about non-humans as ‘what?’, even though this has nothing to do with a default use of the masculine agreement pattern.

6.4.2.5 Tangkic languages

The Tangkic family includes three languages, Lardil (NAa), Kayardild (plus Yangkaal; NAb1) and Yukulta (plus Nguburindi; NAb2), the latter two being more closely related to each other (cf. Evans 1995:9-13).⁵⁶ Of the three languages, only Yukulta and (marginally) Kayardild allow for a ‘who?, what?’ interrogative. Thus, the Yukulta interrogative *ngaka* may be used pronominally as either ‘what?’ (53) or ‘who?’ (54) and attributively as ‘which [N]?’ (55).

Yukulta (non-Pama-Nyungan, Tangkic; NAb2)

- (53) *ngaka = yi = ka = nti* *tiya?*
 IPW[ABS]=2SG.A=TRANSITIVE=FUT eat
 ‘What will you eat?’⁵⁷ (Keen 1983:201)

⁵⁶ Dixon’s (2002) NAc language Minkin appears to be “fairly closely related to the Tangkic languages, though more distant from any of them than they are from each other”, although we have only “scanty nineteenth century materials” for Minkin (Evans 1995:9).

⁵⁷ Keen (1983:219) calls a combination of morphemes such as [*yi + ka + nti*] in this example “the clitic complex [which] is obligatory and is suffixed to the first constituent of the sentence”. This is why the clitic boundary sign = has been added in the Yukulta examples adapted from Keen (1983).

- (54) a. *ngaka = thu = ŋi = ngki ngityintyu[u] wara?*
 IPW[ABS]=1SG.OBL=3SG.S=FUT 1SG.COM come
 ‘Who will come with me?’ (Keen 1983:201)
- b. *ngaka tan-ma tangka?*
 IPW[ABS] this-STATIVE man[ABS]
 ‘Who is this man?’ (Keen 1983:229)
- c. *ngaka-ngaka-ya = lkari kuritya tathin-ta ngawu?*
 IPW-IPW-ERG=3PL.PRS see.IND that-ABS dog[ABS]
 ‘Who are those people looking at that dog?’ (Keen 1983:208)
- (55) *ngaka-ya maku-ya = ka = nta kurka yakuli?*
 IPW-ERG woman-ERG=TRANSITIVE=PST take fish[ABS]
 ‘Which woman took the fish?’ (Keen 1983:200)

In Kayardild, which has a similar interrogative *ngaaka*, the situation is somewhat more complicated, because “when in a core grammatical function,⁵⁸ this normally combines with the [...] four generic nouns” summarized in (56) (Evans 1995:134, 366), so that Evans (1994:206) prefers to treat such phrases as “single lexical item[s]”.

Kayardild (non-Pama-Nyungan, Tangkic; NAb1; Evans 1995:366)

- (56) *ngaaka dangkaa* (+ ‘person’) ‘who?’
ngaaka wuranda (+ ‘food’) ‘which animal/bird/fish/edible plant?’
ngaaka thungalda (+ ‘tree, thing’) ‘what?, which thing/artefact?’
ngaaka nida (+ ‘name’) ‘what name?’

The simple, non-compounded form *ngaaka* is used when the interrogative is marked for “semantic case” (Evans 1995:367), e.g. the genitive *ngaak-arra* ‘whose?’ or the “verbal dative” *ngaaka-maru-tha* ‘for whose benefit?’ vs. the “utilitive” *ngaaka-marra* ‘to use for what?’ or the “privative” *ngaaka-warri* ‘lacking what?, without what?’. Apparently, the human and the non-human interpretation here depend on the semantics of the respective case. Another situation where “the interrogative base suffices” is represented by “no conjecture”-contexts, viz. “where there is insufficient evidence to assign the questioned entity to one of these categories” (Evans 1995:367), as in (57) and (58).

⁵⁸ The “core functions” are “subject, object or indirect object” (Evans 1995:134).

Kayardild (Evans 1995:367)

- (57) *dathin-a dangka-a kiyarrng-ka dangka-a?*
 that-NOM person-NOM two-NOM person-NOM
ngaaka-ntha kurri-n-d
 IPW-ASSOCIATING.OBL look-NMLZ-NOM
 ‘Who/what are those two people looking for?’
- (58) *dan-da ra-yin-da kila ngaak?*
 here-NOM south-FROM-NOM trampling.noise[NOM] IPW[NOM]
 ‘Who/what is trampling from the south here?’⁵⁹

Finally, “the interrogative base suffices [...] in asking someone their identity” (Evans 1995:367), (59).

Kayardild (Evans 1995:367)

- (59) *nyingka ngaak? nyingka ngijin-da ngaak?*
 2SG.NOM IPW[NOM] 2SG.NOM my-NOM IPW[NOM]
 ‘Who are you? You are my what (relation)?’

Note, furthermore, that as an indefinite pronominal *ngaaka* alone also seems to suffice when the intended meaning is human ‘someone’, as in (60a), but not when it is non-human, as in (60b).

Kayardild (Evans 1994:207)

- (60) a. *ngada ngaaka-na kurri-jarra dan-kina bath-ina,*
 1SG.NOM INDF-PRIOR see-PST here-PRIOR west-PRIOR
ngada kinaa-nangku ngumban-ju
 1SG.NOM tell-NEG.POTENTIAL 2SG-FUT
 ‘I saw **someone** here in the west, (but) I won’t tell you (who).’
- b. *ngaaka thungal-da dan-da riin-id?*
 INDF[NOM] thing-NOM here-NOM from.east-STILL
 ‘I saw **something** coming from the east here, what is it?’

It is not clear, however, whether the same preference for the human meaning of the non-compounded form *ngaaka* is found when an indefinitely used *ngaaka* is marked for “semantic case”.

The interrogative pronominals of the third Tangkic language, Lardil, found in the sources are summarized in (61).

⁵⁹ The form *ngaak* is due to the “prosodic truncation of final /a/” (cf. Evans 1995:63-65).

Lardil (non-Pama-Nyungan, Tangkic; NAa; Klokeid 1976:363-364, 369; Capell 1943:49)

- | | | | |
|------|----------------------|----------------------------------|--------------|
| (61) | <i>ngaja-rta</i> | (+ <i>tang(k)a</i> ‘person’) | ‘who?’ |
| | <i>ngaju werne</i> | (+ <i>werne</i> ‘food’) | ‘what?’ |
| | <i>ngaju thungal</i> | (+ <i>thungal</i> ‘tree, thing’) | ‘what?’ |
| | <i>ngajida</i> | (+ <i>nida</i> ‘name’) | ‘what name?’ |
| | <i>ngaji-kan</i> | (+ <i>-kan</i> GEN) | ‘whose?’ |

The first four forms are compounds. However, since *ngaja-rta* is obviously less transparent than the other three, it may need some explanation. The element *-rta* /-ʈa/ in *ngaja-rta* is a truncated form of the noun *tang(k)a* ‘person, man’. Lardil has several other compounds where *tang(k)a* ‘person, man’ is similarly truncated (cf. Klokeid 1976:68-73), as in (62), where the first components are *tiin* ‘this’ and *pata* ‘that (to the west)’. The retroflexion of the initial apico-alveolar *t*- of *tang(k)a* in *ngaja-rta* and *pata-rta* is an absolutely regular phonological process, since in Lardil “word-initial apicals *t*, *n* are retroflexed (apico-domal) following a word ending in a vowel or in a retroflexed consonant” (Klokeid 1976:29).

Lardil (Klokeid 1976:68)

- | | | | | |
|------|-----------------|---------------------|--------------------------|------------------------|
| (62) | NOM | ACC | NON<FUT> | FUT |
| | <i>tiin-ta</i> | <i>tiinin-angan</i> | <i>tiinad-tangkangad</i> | <i>tiinkur-tangkar</i> |
| | <i>pata-rta</i> | <i>patyin-angan</i> | <i>pathad-tangkangad</i> | <i>pathur-tangkar</i> |

According to Klokeid (1976:364), “question phrases take case according to grammatical relations, just like other dependents, and enjoy the same restrictions and freedom of word order”, as in (63a) and (63b). Note that the “question phrase” can be left dislocated, as in (63c), in which case it does not take case, just like any other nominal (cf. Klokeid 1976:366-367).

Lardil (Klokeid 1976:364)

- | | | | | | | |
|------|----|-----------------------------|-------------------|-------------------|-------------|-----------------|
| (63) | a. | <i>nyingki</i> | <i>ngajuwa-n</i> | <i>thungal-in</i> | <i>kudi</i> | <i>kun?</i> |
| | | 2SG.NOM | IPW-ACC | thing.ACC | see | EVENTIVE |
| | b. | <i>ngajuwa-n</i> | <i>thungal-in</i> | <i>kudi</i> | <i>kun</i> | <i>nyingki?</i> |
| | | IPW-ACC | thing-ACC | see | EVENTIVE | 2SG.NOM |
| | c. | <i>ngaju</i> | <i>thungal</i> | <i>nyingki</i> | <i>kudi</i> | <i>kun?</i> |
| | | IPW.NOM | thing.NOM | 2SG.NOM | see | EVENTIVE |
| | | ‘What (thing) did you see?’ | | | | |

It should be pointed out that Klokeid (1976) provides case marked examples only for the compounds *ngaju werne* ‘what?’ and *ngaju thungal* ‘what?’. The interrogative *ngaja-rta* ‘who?’ is always in the nominative, so that it is not clear whether both parts of the compound will take case marking, as e.g. the

structurally similar forms in (62) do. The last compound *ngajida* ‘what name?’ comes from Capell (1943:49).

The genitive form *ngaji-kan* ‘whose?’ in (61) suggests that just like in Kayardild (cf. above), in Lardil the simple, non-compounded form of the interrogative is used when the interrogative is marked for “semantic case”. In Kayardild, the genitive form of the interrogative *ngaak-arra* also means ‘whose?’. Unfortunately, the genitive is the only “semantic case” of the interrogative that can be found in Klokeid (1976), so that it is not clear whether the non-compounded form of the interrogative can also have a non-human meaning.

As can be observed in (61), Lardil appears to have three interrogative pronominal forms differing only in their final vowel: *ngaja*, *ngaju* and *ngaji*. Klokeid (1976:363-364, 375) discusses only the first two of them. In particular, he says that *ngaja* is ‘which (human)?’ and *ngaju* is ‘which (non-human)?’, although the two “are not used freely”. Furthermore, he analyses *ngaja* as derived “from underlying *ngaj* by Augmentation” and *ngaju* as derived from “underlying *ngajuwa* by Apocope and Non-Apical Consonant Truncation”. I find his analysis of *ngaju* as underlyingly *ngajuwa* well justified. At the same time, Klokeid’s analysis of *ngaja* as underlyingly *ngaj* can be improved if the genitive form *ngaji-kan* is taken into account. Thus, given that Klokeid (1976:35-37, 43-45) himself argues that for a disyllabic nominal that in the nominative ends in *-a* the underlying final vowel is the one that shows up in the non-nominative case forms and in the vocative (64), it seems reasonable to analyse *ngaja* as the nominative *ngaj-a* derived from underlying *ngaji*.

Lardil (Klokeid 1976:35-36, 44-45)

(64)	NOM	ACC	NON<FUT>	FUT	VOC	
	<i>thawa</i>	<i>thawa-n</i>	<i>thawa-ngad</i>	<i>thawa-r</i>		‘rat’
	<i>karta</i>	<i>kartu-n</i>	<i>kartu-ngad</i>	<i>kartu-r</i>	<i>kartu!</i>	‘child (of a woman)’
	<i>tunja</i>	<i>tunji-n</i>	<i>tunji-ngad</i>	<i>tunji-wur</i>	<i>tunji!</i>	‘junior wife’s brother’

As to the non-human underlying form *ngajuwa*, it would not be implausible if the vowel *u* would be originally the nominative *-a* that has been assimilated to the following *w*. The final *-wa* of **ngaja-wa* could then be a truncated form of the word *wanka* ‘arm’ (Klokeid 1976:35), for instance, just like *-rta* in *ngaja-rta* is a truncated form of *tang(k)a*. This connection may be further supported by the fact that in Kayardild *wanka-* means ‘branch of tree’ as in *This branch will make a good boomerang* (Evans 1995:780). Note in this respect that the original meaning of the Lardil word *thungal* ‘thing’ used in *ngaju thungal* is ‘tree’. Compare also Kayardild *thungalda*, the cognate of Lardil *thungal*, which is similarly used in the non-human interrogative ‘what?’, *ngaaka thungalda*, and

also has ‘tree, esp[ecially] when seen as source of tools, building materials etc.’ as its first meaning.

Summing up, we can reconstruct for pre-Lardil (i) a single interrogative pronominal stem **ngaji* ‘who?, what?’ (also attributive ‘which [N]?, what [N]?’), which in the nominative had the form **ngaj-a*, as well as (ii) two phrasal interrogative pronominal constructions presumably used in the “non-semantic cases” similarly to Kayardild, the human ‘who?’ **ngaja tangka* ‘what/which person?’ > *ngaja-rta* and the non-human ‘what?’ **ngaja wanka* ‘what/which branch?’ → ‘what?’ > **ngaja-wa* > **ngajuwa* > *ngaju(wa) thungal* ‘what/which tree?’ → ‘what?’.

By way of conclusion, let us compare Lardil with the other Tangkic languages. The original form of the interrogative pronominal root in the Tangkic languages may be brought back easily to something like Lardil *ngaji/ngaj-a*. In Kayardild and Yukulta, the final syllable must have been truncated. The final *-ka* in Kayardild *ngaaka* and Yukulta *ngaka* is likely to be related to the Kayardild and Lardil focus particles *=(a)ka* (Evans 1995:392-393). It may be interesting to point out in this respect that Evans (1995:128) suggests that the Kayardild nominative suffix “may have originated as a discourse particle *|-ka|*”. The tendency to use the interrogative pronominal in phrasal constructions with generic nouns in the Tangkic languages can be compared to the use of gender markers on interrogative pronominal roots in many other non-Pama-Nyungan languages (cf. Section III.6.3.3). Given that the interrogative **ngaji* (and later also **ngaj-a*) ‘who?, what?’ (also attributive ‘which [N]?, what [N]?’) appears to go back to Proto Tangkic, any hypothesis on the origins of this lack of differentiation is very likely to involve some other non-Pama-Nyungan languages. The neighbouring Garrwan languages do not show immediate formal parallels to the Tangkic interrogative, even though Proto Garrwan also appears to have had a ‘who?, what?’ interrogative, just as modern Garrwa does (Section III.6.4.2.6). Among Gunwinyguan and other Arnhem Land languages, one often encounters locative(-cum-selective) interrogatives based on the form *(n)gaya* or the like (cf. Section III.6.3.2), which shows more similarity to the Tangkic form. In this respect, it may be interesting to mention that according to Evans (1995:38-39), the speakers of Proto Tangkic are very likely to have “originally liv[ed] just to the south of the Arnhem Land escarpment, possibly along the upper reaches of the Roper River”, i.e. “adjacent to the non-Pama-Nyungan languages of Arnhem Land to which they are most closely related genetically”.

6.4.2.6 Garrwan languages

The Garrwan language family consists of Wanyi (X1) and Garrwa (X2). Garrwan languages have sometimes been classified as Pama-Nyungan, but now they are usually regarded as non-Pama-Nyungan. Evans (2003a:10) argues that of the

non-Pama-Nyungan languages, Proto Garrwan may be the closest relative of Proto Pama-Nyungan. In what follows, I will first present the Garrwan interrogatives as they are found in the modern languages and I will conclude this section by a discussion of their possible etymologies.

Garrwa is further subdivided into an Eastern and a Western variety. Eastern Garrwa has a single interrogative ‘who?, what?’ and a different interrogative ‘where?’, while in Western Garrwa “the same [two] roots [...] are used more or less interchangeably” (Breen 2003:441). Wanyi is different in having dedicated ‘who?’, ‘what?’ and ‘where?’ interrogatives. The inflected forms of the respective interrogatives in the three Garrwan idioms are summarized in Table 5. The empty Wanyi and Eastern Garrwa cells appear to be due to the lack of data (cf. Breen 2003:426-427, 440). The only empty Western Garrwa cell, in the ablative row, means that this form does not exist (cf. Furby 1972:18). Since the paradigms of the interrogatives in Table 5 demonstrate a good deal of irregularities as compared to the ordinary nominal system (cf. Breen 2003:434-436), I found it useful to highlight these irregularities in bold. The irregularities in the paradigms of Garrwan interrogatives are of two kinds. The first kind of irregularity concerns the roots of the interrogatives. The initial *w-* may be dropped in some forms and *a* may sometimes be replaced by *i* and in one case by *u*. Second, the interrogatives differ from nouns in their case endings, even though in most instances the endings are not completely unrelated. There is considerably more similarity with the inflections of demonstratives and personal pronouns (cf. Breen 2003:438-441; Furby 1972). For instance, in Western Garrwa the ergative marker on most nouns is *-wanyĩ*, the second allomorph, *-nyĩ*, is used on adjectives ending in *rra*, on singular possessive pronouns, and after the dual, plural, concomitant and ‘deceased’ suffixes (Breen 2003:434). At the same time, the ergative is marked with *-ngi* on the demonstrative ‘that’, *nana-ngi*, and with *-ni* on ‘this’, *nangi-ni* or *nanangi-ni* (Furby 1972:15-16), which is already much closer to the ergative suffixes found on the interrogatives.⁶⁰ Yet another example could be the initial *n* in the Western Garrwa plural suffix *-(n)muku*. This *n* appears when the suffix is used with demonstratives and interrogatives, but not with nouns (cf. Section III.6.3.4 for another example). Recognizing the similarities, it should be kept in mind, though, that the inflectional paradigms of

⁶⁰ As can be readily observed, should the morpheme boundary in the proximate demonstrative be drawn before *ngi* instead of *ni*, i.e. **na-ngini/*nana-ngini*, we would have got an even better parallel with the ergative interrogatives: *wanyja-ngini* vs. **na-ngini/*nana-ngini*, on the one hand, and *wanyi-ngi* vs. **nana-ngi*, on the other. However, even though diachronically this analysis is quite plausible, given the existence of such allomorphs of the proximate demonstrative stem as *nayĩ* and *na-* (Furby 1972:15), synchronically it is somewhat difficult to maintain. For instance, the plural suffix *-nmuku* is added before the case endings and the proximate demonstrative root is then *nangi-* (Furby 1972:15).

Table 5. Garrwan interrogatives ‘who?’, ‘what?’, ‘where?’ and ‘where?’ (based on Breen 2003:441)

	Wanyi ‘who?’	Wanyi ‘what?’	Wanyi ‘where?’	Eastern Garrwa ‘who?, what?’	Eastern Garrwa ‘where?’	Western Garrwa ‘who?, what?, where?’
ABS	(w)inyjika	wanyi	winyja	wanyi	wanyja	wanyi
ERG ⁱ⁾	winyjika-ni	wanyi-ni	—	wanyi-ni	—	wanyi-ngi
DAT ⁱⁱ⁾	(w)inyjika-nyi, winyjika-aninyi	wanyi-ngkanyi	—	wanyi-ngkanyi	—	wanyi-ngkanyi
LOC	—	—	winyja-na	—	inyja-ni, inyja-wa, wanyja-wa	wanyi-na
ALL	—	—	wunyju-ku	—	—	wanyi-ngkurri
ABL	—	—	winyji-bunyi, inyja-banyi	—	wanyja-bunanyja, wanyja-bananyi	wanyi-biyurri
DU[ABS]	—	—	—	—	—	—
PL[ABS]	—	—	—	—	wanyi-muku ⁱⁱⁱ⁾	wanyi-nkuya, wanyja-nkujarra wanyja-nmuku

- i) In Wanyi, ergative is also instrumental. Breen (2003) calls this combination “operative”. In Garrwa, instrumental is usually expressed with locative. Both in Wanyi and Garrwa, “instrumental function may also be fulfilled by concomitant (having) plus [...] ergative” (Breen 2003:435).
- ii) At least in Western Garrwa, dative “indicates benefaction and possession” (cf. Furby 1972:30).
- iii) sic! Shouldn’t this form be in the preceding column?

The irregularities in the paradigms of Garrwan interrogatives are of two kinds. The first kind of irregularity concerns the roots of the interrogatives. The initial *w-* may be dropped in some forms and *a* may sometimes be replaced by *i* and in one case by *u*. Second, the interrogatives differ from nouns in their case endings, even though in most instances the endings are not completely unrelated. There is considerably more similarity with the inflections of demonstratives and personal pronouns (cf. Breen 2003:438-441; Furby 1972). For instance, in Western Garrwa the ergative marker on most nouns is *-wanyĩ*, the second allomorph, *-nyĩ*, is used on adjectives ending in *rra*, on singular possessive pronouns, and after the dual, plural, concomitant and ‘deceased’ suffixes (Breen 2003:434). At the same time, the ergative is marked with *-ngi* on the demonstrative ‘that’, *nana-ngi*, and with *-ni* on ‘this’, *nangi-ni* or *nanangi-ni* (Furby 1972:15-16), which is already much closer to the ergative suffixes found on the interrogatives.⁶¹ Yet another example could be the initial *n* in the Western Garrwa plural suffix *-(n)muku*. This *n* appears when the suffix is used with demonstratives and interrogatives, but not with nouns (cf. Section III.6.3.4 for another example). Recognizing the similarities, it should be kept in mind, though, that the inflectional paradigms of demonstratives and personal pronouns are themselves quite irregular, with an important amount of lexically conditioned allomorphy.

Let us now discuss the Garrwan interrogatives from an etymological perspective. To begin with, in all probability the Wanyi interrogative *winyjika* ‘who?’ has been derived from *winyja* ‘where?’, as is not uncommon in Australia and, for instance, is also found in the neighbouring Pama-Nyungan language Wakaya, where *winthi-nga* ‘who?’ “is clearly derived from *wiinthi* ‘where’ with the ‘after’ or ‘out of’ formative *-nga*” (Breen 1974:3.3; cf. Section III.6.3.2). The final *-ka* of Wanyi *winyjika* may be compared, among other things, to the element *-ka* in the neighbouring Yukulta *ngaka* ‘who?, what?’, discussed in Section III.6.4.2.5, or the Garrwa “reflexive” pronominal suffix *-ngka/-ka* (cf. Furby 1972:2-3), which e.g. in (65) does not really look like a prototypical reflexive.

⁶¹ As can be readily observed, should the morpheme boundary in the proximate demonstrative be drawn before *ngi* instead of *ni*, i.e. **na-ngini/*nana-ngini*, we would have got an even better parallel with the ergative interrogatives: *wanyja-ngini* vs. **na-ngini/*nana-ngini*, on the one hand, and *wanyi-ngi* vs. **nana-ngi*, on the other. However, even though diachronically this analysis is quite plausible, given the existence of such allomorphs of the proximate demonstrative stem as *nayĩ*- and *na-* (Furby 1972:15), synchronically it is somewhat difficult to maintain. For instance, the plural suffix *-nmuku* is added before the case endings and the proximate demonstrative root is then *nangi-* (Furby 1972:15).

Wanyi (non-Pama-Nyungan, Garrwan; X1; Furby 1972:3)

- (65) *jabulaba na-ngka ba(r)n-kanyi*
 hungry 3SG-REFL meat-DAT
 ‘He is hungry for meat.’

The ‘where?’ interrogatives *winyja* of Wanyi, *wanyja* of Eastern Garrwa, and the cognate ‘who?, what?, where?’ root *wanyja-* of Western Garrwa are identical to the common Pama-Nyungan ‘where?’ interrogative *wanhdha/wanyja*, with “various other forms that show assimilations” (cf. Dixon 2002:332-334; see Section III.6.3.2). Breen (2003:441) also suggests a parallel with *inyja* ‘which?’ and *inyja-ni* ‘where (to)?’ (*-ni* is LOC/ERG) in neighbouring Wambaya (Mirndi group; Nordlinger 1993:4.7.2). Both the Garrwan and the Wambaya interrogatives at issue have probably been borrowed from Pama-Nyungan. The direction of borrowing is suggested by (i) the widespread occurrence of *wanhdha/wanyja* in Pama-Nyungan, (ii) by the fact that this interrogative is more readily analyzable from the point of view of Pama-Nyungan morphology than from that of Garrwan or Mirndi languages, (iii) by the fact that in Wambaya the locative suffix *-ni* needs to be added to *inyja* for it to be able to function as ‘where?’ (compare also the Western Garrwa absolutive form *wanyja-ni*, while as a rule absolutive is unmarked in this language).

The third Garrwan interrogative, *wanyi*, may contain some frozen non-absolutive case suffix, such as *-(wa)nyi*. Thus, *-wanyi* is the usual ergative allomorph in Garrwa. The form *-nyi* is identical to one allomorph of the ergative and dative suffixes in Garrwa and Wanyi, it also makes part of most other allomorphs of the dative (*-anyi*, *-ngkanyi*, *-yngkanyi* other allomorphs are *-yngka* and *-wa*) and of one allomorph of the locative in Garrwa (*-nyina*, other allomorphs being *-ina* and *-na*; cf. Breen 2003:434). In the languages of the neighbouring Mirndi group, one finds a non-absolutive masculine affix *-ni/nyi-* (reconstructed by Harvey, Green & Nordlinger 2006:308 as **ni-*), where non-absolutive minimally implies ergative, but usually ergative/locative/instrumental.⁶² The use of the ergative allomorph *-ni* in Wanyi *wanyi-ni* may suggest that *-nyi* used to be a dative or at least it used to be interpreted as dative (which would also include benefactor and possessor marking), because in Wanyi *-ni* is the ergative allomorph used “after DAT suffix” (cf. Breen 2003:434).

⁶² Thus, Wambaya, Jaminjung and Ngaliwurru have the ergative/locative/instrumental suffix *-ni*, while Nungali usually marks the ergative/locative with either the prefix *nyi-* or the suffix *-ni* (Nordlinger 1993:4.4.1.3). Jingulu has a non-feminine ergative suffix *-(r)ni* (it is also mostly animate, since “inanimate transitive subjects generally take the Instrumental suffix */-arndi/*”, cf. Pensalfini 1997:273).

As to the original semantics of *wanyi*, I disagree with Dixon (2002:328-329), who suggests that the original meaning of *wanyi* “‘what’ [was] extended [in Garrwa] also to have the sense ‘who’”. I believe that *wanyi* was used both as ‘who?’ and ‘what?’ and probably also as ‘where?’, the locative meaning presumably being the oldest of the three.

To begin with, the development of a ‘who?, what?’ interrogative from a ‘what?’ interrogative by a purely semantic shift would be extremely difficult to account for semantically. On the contrary, one could easily imagine a general interrogative meaning ‘who?, what?’ to become restricted to the meaning ‘what?’, when a separate human interrogative ‘who?’ evolves. Indeed, as discussed above, Wanyi appears to have recently formed a dedicated ‘who?’ interrogative *winyjika* ‘who?’ from its locative interrogative *winyja* ‘where?’.

Second, that the meaning ‘where?’ of *wanyi* in Western Garrwa is unlikely to be a recent innovation is suggested by the fact that *wanyi* may have this meaning next to ‘who?’ and ‘what?’ without any additional locative morphology. In turn, that ‘where?’ may be the oldest meaning of *wanyi* would square well with the assumption that *wanyi* contains a frozen non-absolutive suffix *-(wa)nyi*.⁶³ As has been mentioned, the label non-absolutive usually implies at least ergative and locative and/or instrumental. Of these three functions, “the locative function seems more semantically basic”, in the sense that extensions from locative to other functions are “widely attested” and are semantically more easy to conceive than the reverse change (cf. Nordlinger 1993:4.4.1.3). A further extension from ‘where?’ to ‘which (one)? (person or thing)’ and ultimately to ‘who?, what?’ is semantically quite plausible as well. The latter process may have been further stipulated by the presumed borrowing of *wanyja* ‘where?’ from some Pama-Nyungan source.

6.4.2.7 Rembarrnga

Rembarrnga is a non-Pama-Nyungan (Eastern) Gunwinyguan language (NBc1). Rembarrnga interrogative pronominals and some of the related interrogative pro-forms are summarized in (66).

⁶³The Garrwan interrogative *wanyi* may be cognate to the bound interrogative root *(w)ani-* of Jingulu, a geographically close language of the Mirndi group (cf. Section III.6.4.2.3). In Jingulu, *(w)ani-* occurs in the following compound interrogatives as *(w)ani.kiy-* ‘do what?’, *(w)ani.ngkiliji* ‘how many?, how much?’ and *(w)ani.kirrkiji* ‘what sort of [N]?’ (Pensalfini 2003:146-147). The use of ‘where?’, which could presumably be the original meaning of *(w)ani*, as ‘how?’, ‘(do) what?’ and ‘which/what [N]?’ would not be unusual in the Australian context.

Rembarrnga (non-Pama-Nyungan, Gunwinyguan; NBc1; following Adam Saulwick, p.c.)

- | | | |
|------|--------------------|-------------------------------|
| (66) | <i>yana?</i> | ‘what?’, maybe also ‘who?’ |
| | <i>yana?-gurn</i> | ‘who?’ |
| | <i>yana?-gan</i> | ‘why?, what for?’ (dative) |
| | <i>yene?</i> | ‘what?, how?’ |
| | <i>yene?-gan</i> | ‘why?, what for?’ (dative) |
| | <i>yene?-ja(m)</i> | ‘where?’ (locative) |
| | <i>yene?-wala</i> | ‘from where?’ (ablative) |
| | <i>yene?-ga?</i> | ‘towards where?’ (allative) |
| | <i>yene?-ji</i> | ‘when?’ (“temporal locative”) |

McKay (1975:114) claims that *yana?* “may not refer to humans”, i.e. it means ‘what?’. In contrast to *yene?* ‘what?’ referring to “space, time or activity”, *yana?* ‘what?’ refers to “specific objects”. However, Adam Saulwick (p.c.) found no restriction on the use of *yana?* in questions about humans in his own corpus. He suggested examples (67a) and (68), also used in his (2003) dictionary, to illustrate the use of *yana?* as ‘who?’. As to McKay’s claim about more concrete reference of *yana?* as compared to *yene?* (in questions about non-humans), Adam Saulwick (p.c.) suggested that examples (69a) and (70) may be viewed as supporting it.

Rembarrnga (Adam Saulwick, p.c.)

- (67) a. *yana? danda-ma?*
 IPW[ABS] 2M-PTCL
 ‘[A:] Who are you?’
- b. *yana?gurn danda?*
 IPW[ABS] 2M
 ‘[A:] Who are you?’
- c. *nginda wamut nga-bol?-miny-gœ*
 1M PROP[ABS] 1-come-PST.PFV-2DAT
 ‘[B:] I’m wamut. I came for you.’
- (68) *bi yana?-yi? ga-gu?-berde?-ra?*
 man[ABS] IPW-ERG NON<PST>.3>3-body-carry.on.shoulder-FUT
 ‘Who’s going to carry the body?’
- (69) a. *yana? nginy-nyawk?*
 IPW[ABS] 2-speak[NON<PST>]
 ‘[A:] What do you speak?’
- b. *nga-nyawk rembarrnga*
 1-speak[NON<PST>] Rembarrnga[ABS]
 [B:] I speak Rembarrnga.’

- (70) *yene?* *bak-nyawk-moern*
 IPW[ABS] [3]-BEN-speak-PST.CONTINUOUS
danguny yene? *bak-ngeji?*
 story[ABS] IPW[ABS] [3]-BEN-tell.about[NON<PST>]
 ‘What was he telling him, what story was he telling?’

It seems, however, that *yana?* in (67a) can and in all probability should be interpreted not as a human interrogative ‘who?’ but as ‘what?’, just as McKay (1975) did. Note that (67a) looks very much like a KIND-question with ‘what?’-dominance (cf. Section II.2). This is suggested by the answer in (67c), where *wamut* is a name of a male “subsection” within the Rembarrnga kinship system. The term “subsection” refers to a group of individuals, therefore the question in (67a) is probably best translated as ‘What are you? (i.e., ‘What is your position in our (Rembarrnga) kinship system?’) and consequently, the answer in (67c) should better involve an indefinite article, ‘I’m *a* wamut’. Note, in this respect, that in the original translation the word *wamut* is written with a small letter, i.e. as a common noun rather than a personal proper name.

Further support for this interpretation comes from another Gunwinyguan language Nunggubuyu (NBd2), as described by Heath (1984). Thus, somewhat similarly to Rembarrnga, Nunggubuyu has *yaji* ‘what?’ and *yaji-ɲuŋ* ‘who?’⁶⁴ and “occasionally, *yaji* may be used with human reference when the focus is not on an individual’s personal identity, but rather on clan or moiety status or the like (cf. English ‘What are you, a Tory or a Laborite?’)”, even though “specifically human forms are more common (in Nunggubuyu [apparently, just like in Rembarrnga, (67b)]) even in such contexts” (Heath 1984:456).

As to example (68), *yana?* appears to be simply used attributively with a human noun *bi* ‘man’, i.e. *bi yana?-yi?* is ‘what man? (ERG)’,⁶⁵ just like *yene?* in *danguny yene?* ‘what story? (ABS)’ in (70).

6.5 Concluding remarks

In this final section I will do two things. First, I will comment on some diachronic and synchronic generalizations about the interrogatives ‘who?’, ‘what?’ and ‘who?, what?’ in Australian languages made by Dixon (2002). I believe that in the light of the discussion in Sections III.6.3-III.6.4 these generalizations require substantial revision. Second, I would like to take a global

⁶⁴ In Nunggubuyu, *-ɲuŋ* is the human singular suffix (Heath 1984:457). The element *-gurn* /-gun/ Rembarrnga *yana?-gurn* ‘who?’ is not immediately analyzable (Adam Saulwick, p.c.).

⁶⁵ Note that *bi* ‘man’ in (68) cannot refer to the body to be carried. This is so because of its meaning and the fact that ‘body’ is already expressed by *gu?* incorporated in the verb *ga-gu?-berde?-ra*.

look at the Pama-Nyungan facts discussed in Section III.6.4.1 and the findings from the remaining non-Pama-Nyungan part of the continent discussed in Section III.6.4.2. The thing I am most interested in is how different or similar the two blocks turned out to be with respect to the lack of differentiation between ‘who?’ and ‘what?’.

6.5.1 On Dixon’s (2002) generalizations

First, Dixon (2002:328) claims that “in most of [the languages with ‘who?, what?’ interrogatives] (at least) the neutralisation [of ‘who?’ and ‘what?’] has been a recent, language-particular change”, and what is more, “in almost all instances it appears that ‘who’ has been extended also to cover ‘what’”. However, as argued in Section III.6.4, in most cases ‘who?, what?’ interrogatives appear to be old enough to be reconstructed for the respective proto languages and one should rather speak of specialization in the languages which now have dedicated ‘who?’ and ‘what?’ interrogatives.

Second, Dixon (2002:329) notes that “from the grammatical descriptions available there does seem to be a TENDENCY for languages that have one form for both ‘who’ and ‘what’ to assign this an exclusively interrogative meaning”. Thus, “for most of the languages in which one form covers both ‘who’ and ‘what’, the available grammars do not give any additional indefinite sense”, with only “one known counter-example [the Tangkic language] Kayardild [(NAb1), where] *ngaaka* covers both ‘who’ and ‘what’ and also has indefinite sense ‘someone’ and ‘something’”. Given that usually in Australian languages interrogatives and indefinites have the same form or at least are clearly morphologically related, such a tendency could have been important. However, as has already been pointed out in Section III.6.3.1, “counter-examples” appear to be much more numerous than is assumed by Dixon. Thus, in at least 5 of the 16 “languages” that have proven to allow for a lack of differentiation between ‘who?’ and ‘what?’, the form used as the interrogative ‘who?, what?’ can also be used both as a human and a non-human indefinite pronominal, either on its own or with some additional morphology. In at least four more idioms (including Kayardild) indefinites are, in principle, also related/identical to the respective interrogatives, but it is not completely clear whether in the indefinite use they can also refer to both humans and non-humans. Furthermore, one more language, viz. Jingulu (Section III.6.4.2.3), may also belong here, provided it really has a ‘who?, what?’ interrogative. For most of the remaining languages, I simply do not have the relevant information.

Third, to explain “why a form which is ‘who’ in many languages may also turn up as ‘what’ in a few (and vice versa although this appears to happen to a lesser degree)”, there usually appears to be no need to resort to such particular scenarios as “taboo[ing of the earlier word for ‘what?’] due to similarity with the

name of some person who died” and its subsequent replacement by “the S form of ‘who?’” (Dixon 2002:331). To this it should be added that a sheer replacement of ‘what?’ by ‘who?’ (or vice versa) is hardly tenable semantically.

6.5.2 *The lack of differentiation between ‘who?’ and ‘what?’: Pama-Nyungan vs. non-Pama-Nyungan*

The Pama-Nyungan languages with ‘who?, what?’ interrogatives appear to be much more homogeneous in the forms and origins of these interrogatives than the non-Pama-Nyungan languages. In almost all Pama-Nyungan languages with such interrogatives, they can be demonstrated to go back to an earlier form *ngana* (or the like) ‘who?, what?’. In the modern Pama-Nyungan languages, the form *ngana* (and related forms) is also by far the most common interrogative pronominal form (cf. Dixon 2002:334, who prefers the form *nga(:)n-*, however). Furthermore, throughout Pama-Nyungan *ngana* and related forms can be found as ‘who?’, ‘what?’ and ‘who?, what?’, which is yet another piece of evidence suggesting that **ngana* should be reconstructed in the latter general ‘who?, what?’ meaning for Proto Pama-Nyungan. Another interrogative, which is equally widespread throughout the Pama-Nyungan area, but somewhat more consistent semantically seems to be the locative interrogative ‘where?’, which should probably be reconstructed as **wa* (cf. an overview in Dixon 2002:332-334, who prefers the form *wanh-*).⁶⁶

In having just two major interrogatives, one general pronominal ‘who?, what?’ and one locative ‘where?’, Proto Pama-Nyungan would be similar to several non-Pama-Nyungan families and languages discussed in Section III.6.4.2, e.g. Nyulnyulan (Section III.6.4.2.4), Tangkic (Section III.6.4.2.5) and Eastern Garrwa of the Garrwan family (Section III.6.4.2.6). It is interesting to note in this respect that (i) Proto Garrwan and Proto Tangkic are sometimes claimed to be the closest relatives (in the above order) of Proto Pama-Nyungan (cf. Figure 1 in Section III.6.1), (ii) Nyulnyulan languages show some “intriguing [...] morphological correlations” with the languages of Mirndi group (Stokes & McGregor 2003:61), which might suggest that Proto Nyulnyulan was spoken more to the (north-)east of where the modern languages are spoken, (iii) the discontinuous Mirndi language group is traditionally spoken in the vicinity of the presumed Proto Tangkic homeland, more precisely to the west and to the south(-east) of it (Evans 1995:38-39; cf. Section III.6.4.2.5), (iv) most non-Pama-Nyungan families with a lack of differentiation between ‘who?’ and ‘what?’ happen to border onto the Pama-Nyungan area (cf. Map 4 in Section III.6.2). Furthermore, many other non-Pama-Nyungan languages distinguishing their

⁶⁶ As has already been discussed elsewhere (cf. Sections III.6.3.4, III.6.4.1.1), the *-nh*, or rather *-nha*, of *wa-nh(a)-* is most likely to be the frozen specificity/focus marker **-nha*.

interrogatives ‘who?’ and ‘what?’ only through the use of gender markers on an interrogative pronominal root as such indifferent to the human vs. non-human distinction should probably be considered here as well (cf. Section III.6.3.3).

It is difficult to say with certainty what the ultimate source of these ‘who?, what?’ interrogatives in the respective proto languages may have been. Some hypotheses concerning the particular families have already been advanced in the respective subsections of Section III.6.4.2. Here, I will attempt to make a broader generalization, taking into account the Pama-Nyungan facts. Admittedly, what follows is rather speculative.

It would probably be most reasonable to take as the starting point for any hypothesis about such a deep linguistic prehistory the common functional and formal patterns of the interrogative pronominals that are readily observable in modern Australian languages, as presented in Section III.6.3. Besides, one should also look for any recurrent similarities in the forms of both non-selective and selective interrogative pronominals, as well as locative interrogatives, since the latter tend to develop into interrogative pronominals in Australian languages (cf. Section III.6.3.2). Thus, similar to the presumed Proto Pama-Nyungan form **ngana* (or the like) a good deal of non-Pama-Nyungan pronominal and locative interrogative roots appear to be based on the velar segment *(n)g(a)*,⁶⁷ e.g. Tangkic **ngaji* ‘who?, what?’, Nyulnyulan **yangki*, Burarra *yina* (AG)-*ga-ya* ‘where?’, AG-(*yi*)*n.ga* ‘who?, what?’, AG-*an.ga-ya* ‘which one?’ (*-ya* is the “realis” marker), Worrorra *ang(k)u*-AG ‘who?, what?’, as well as various (mostly locative-cum-selective) *ga/ka*-based interrogatives mentioned in Section III.6.3.2. The nasal, /n/ or /ŋ/,⁶⁸ attested in many of the forms may reflect the frozen masculine gender marker, just like *an-* in Burarra AG-*an.ga-ya* ‘which one?’ (cf. Glasgow 1984:43).⁶⁹ The form *nV* is for instance so common in non-Pama-Nyungan languages as the masculine gender marker that Harvey (2003:499-500) suggests to reconstruct **na-* as the masculine gender marker for most non-Pama-Nyungan languages. Recall in this respect that in quite a few non-Pama-Nyungan languages the masculine (i) is the default agreement pattern on certain targets and

⁶⁷ The sequence *(n)g* here stands for /g, ng, ŋg, ŋ/, as well as /k, nk, ŋk/.

⁶⁸ Presumably, the velar nasal results from an assimilation or fusion with the following velar stop.

⁶⁹ A different source for the nasal may be suggested by the Burarra non-selective interrogative pronominal root AG-(*yi*)*n.ga* ‘who?, what?’ (non-Pama-Nyungan, Maningrida; NBf1; Glasgow 1984:43), and perhaps the Wardaman interrogative *yinggiya* ‘who?’ (non-Pama-Nyungan, Wagiman-Wardaman; NB12; Merlan 1994:153-157), which may be related. At least the Burarra interrogative appears to be derived from the locative interrogative construction *yina* (AG)-*ga-ya* ‘where?’ (cf. the following footnote) through the omission of the realis suffix *-ya*, contraction of the rest and addition of the agreement prefix. The omission of the realis suffix squares well with the shift from a predicative to a substantive function.

in certain contexts, (ii) often it gets frozen on some elements, and (iii) often it is the only agreement pattern that leaves clear traces when the agreement morphology is lost (or moved to another position in the wordform, e.g. from the prefixed to suffixed position; cf. Sections III.6.3.3-III.6.3.4, III.6.4.2.2.2). Could not, therefore, the final segment *-na* of the Proto Pama-Nyungan **ngana* ‘who?, what?’ be a reflex of such a masculine marker as well? As far as the non-nasal element *ga* is concerned, the existence of forms like Maranunggu *ka* ‘(be) where?’ (non-Pama-Nyungan, Western Daly; NHb1; Tryon 1970:71), the frequent occurrence of locative interrogatives among the aforementioned interrogatives, and the common link between ‘where?’ and (non-)selective interrogative pronominals may suggest that this element *ga* used to be the locative interrogative ‘(be) where?’.⁷⁰

⁷⁰ Actually, this *ga/ka* locative interrogative might also be related to similar looking demonstratives in various Australian languages and the 3MIN non-past prefix **ka-* reconstructed by Harvey (2003:499) for the majority of the non-Pama-Nyungan languages. A possibility of such a link may be suggested by the Burarra data in Glasgow (1984). Thus, in Burarra the locative interrogative must be constructed as *yina* (AG)-*ga-ya*. Besides the gender agreement marker AG, its consists of (i) the “realis” suffix *-ya* (1984:18-19), as in *gun-guni-ya* ‘this is the one, now’ (from *gun-guna* ‘G4-this’; 1984:4), (ii) the element *ga*, glossed by Glasgow (1984:18) as “place”, and (iii) the “mood word *yina* ‘interrogative’” (1984:41-43). The element *ga* “place” is analyzed by Glasgow (1984:18-20) as the distal demonstrative root, as in *ga-narda* ‘there near you’ vs. *ngunyu-narda* ‘here near you’ (1984:18) or *gun-ga-ya ga-ta* ‘the one placed there’, lit. G4-“place”-REAL “place”-in.sight (1984:20). In the ‘where?’-construction the element *yina* is normally obligatory. However, when the construction is followed by *wenga* ‘from’, *yina* can be replaced by the gender G4 oblique prefix *gu-*, as in ‘where from?’ *gu-ga-ya wenga* vs. *yina ga-ya wenga* (cf. 1984:19, 43). Gender G4 includes “wood (generic [...]), water, fire, places, houses, furniture” (1984:8; italics added). One may hypothesize then that in *yina* (AG)-*ga-ya*, the “mood word *yina* ‘interrogative’” is the original locative interrogative, while *ga* is just a deictic ‘(be) there’, which apparently is being reanalyzed as the locative interrogative proper. That is, originally *yina* (AG)-*ga-ya* literally may have meant something like ‘where (is it that) s/he/it is there?’. This hypothesis is further supported by the fact that as the “interrogative” word, *yina* appears to be used only in combination with *ga* (cf. 1984:41-43). Furthermore, one finds interrogatives similar to *yina* in many non-Pama-Nyungan languages, and usually with locative or circumstantial semantics, as in Warrwa *jana* ‘where?, (do) what?’ (Nyulnyulan; NE1; McGregor 1984), Mangarayi *jana* ‘where?’ (Maran or Gunwinyguan; NBa; Merlan 1982), Gaagudju *yaa-na* ‘where?, which?’ (*-na* is G1 suffix, cf. Harvey 2002:195-196) and *ya*-AG-‘be, do, say’ ‘(be, do, say) how/what/when?’ (NBk; Harvey 2002), Kayardild *jina* ‘where?’ (Tangic; NAb1; Evans 1995), Gooniyandi *yiniga* ‘how?’ (Bunaban; NF2; McGregor 1990), Rembarrnga *yene?* ‘how?, where?, what?’ and *yana?* ‘what?’ (Gunwinyguan; NBc1; Adam Saulwick, p.c.), Ngalakan *yana?* ‘what?, do what?’ (Gunwinyguan; NBc2; Merlan 1983).

Summing up, the relatively frequent occurrence in Australia of general interrogative pronominal roots and words as such not differentiating between human and non-human meanings appears to be attributable to the following recurrent tendencies: (i) the tendency for the locative interrogative ‘where?’ (particularly, ‘be where?’) to have a parallel selective usage as ‘which one?’ and to further develop non-selective interrogative pronominal usage, (ii) the tendency for the locative/(non-)selective interrogatives to be marked for gender, (iii) the tendency for the masculine interrogative pronominal to be the default form in questions about humans and thus the closest equivalent of ‘who?’, accompanied by a tendency for the default use of the masculine agreement pattern, irrespectively of the human/non-human opposition, (iv) the tendency to loose gender, with masculine forms tending to be the only ones to survive due to the default use of the masculine agreement pattern.

7 North America

7.1 Introduction

North America is conventionally defined here as the region stretching from the USA and Canada in the north to Panama in the south, with the islands of the Caribbean belonging to South America. For the reasons explained in Section III.1, it happened to be impossible to examine the languages of North America in the same degree of detail as many of the Old World languages in the previous sections. Still, I have found several North American languages that appear to allow for the lack of differentiation between ‘who?’ and ‘what?’.¹ The northernmost of these languages belong to the Algonquian branch of the Algic family and are spoken in the north of the USA and the south of Canada. The Algonquian language considered in more detail in Section III.7.2, Nishnaabemwin, belongs to the Ojibwa dialect continuum. In Section III.7.3, I present the data of Timucua, an extinct language of Georgia and Florida which is currently considered as an isolate, but has also been argued to be related to some South American languages. Of all the North American languages discussed here, Timucua is probably the least certain case, as far as the presumed lack of differentiation between ‘who?’ and ‘what?’ is concerned. The following North American language discussed in Section III.7.4 is Timbisha/Panamint, a moribund Uto-Aztecan language of southeastern California and southwestern Nevada. In Section III.7.5, I examine several Mayan languages spoken primarily in Guatemala and the neighbouring regions of Mexico. Finally, in Section III.7.6, two closely related Arawakan languages, Garifuna/Central American Carib and (Dominican) Island Carib, are discussed. In fact, the two languages are Caribbean and ultimately South American in origin. Still, I present them here because Garifuna is currently spoken in Belize, Guatemala, Honduras and Nicaragua, and Dominican Island Carib, by now extinct, was its closest relative. All the other Arawakan languages are discussed together with the other South American languages in Section III.8.2.

The presumed lack of differentiation between ‘who?’ and ‘what?’ in the North American languages discussed may have various sources. Thus, at least the following possibilities seem to be attested: (i) the peculiarities of the gender semantics (in the Algonquian languages), (ii) interference from the dominant vernacular (possibly Algonquian again), (iii) the development of the locative ‘where?’ interrogatives into selective and subsequently non-selective

¹ Note that the North American French varieties and derived creoles allowing for a lack of differentiation between ‘who?’ and ‘what?’, such as Louisiana French and Louisiana Creole, are discussed elsewhere (cf. Section III.3.1.3.2).

interrogative pronominals (Timbisha, probably one Mayan language, and tentatively, Timucua), (iv) the origin of the respective interrogatives in clausal constructions meaning something like ‘the one that [P] is called?...', implying ‘the one that [P] is who/what/which one?’ (most Mayan languages discussed).

7.2 Nishnaabemwin (Ojibwa)

Nishnaabemwin, as described by Valentine (2001), is an Algonquian idiom spoken in Canada on the shores of Lake Huron and Georgian Bay, and inland points to the east. It belongs to the so-called Ojibwa or Anishinaabemowin dialect continuum (cf. Valentine 2001:1) and itself comprises two major varieties Odawa and Chippewa/Eastern Ojibwe.² The Ojibwa dialect continuum belongs to the Central Algonquian languages, which together with the Plains and Eastern Algonquian languages constitute the Algonquian branch of the Algic family. Algonquian languages are traditionally spoken on a vast territory including most of eastern and southern Canada and the northeastern quarter of the USA. The two non-Algonquian Algic languages, Wiyot and Yurok, are spoken in a small area in the north of California.

Like the other Algonquian languages, Nishnaabemwin distinguishes two genders, animate and inanimate. The same distinction is also made in the interrogative pronominals, viz. animate *wene(n)/wenesh* and inanimate *wegne(n)/wegnesh/wanesh/ogonen* (Valentine 2001:129, 979-981).³ The two pronominals can also be used attributively as ‘which [N]?’, normally agreeing in gender with the head noun (Valentine 2001:981). Examples (1-3) illustrate the use of the two interrogative pronominals.

Nishnaabemwin

- (1) *wenesh* *gaa-daapnang* *nmookmaanens?*
 IPW.AN 3SG.AN.PROX>3INAN.took my.knife.DIM[INAN.SG]
 ‘Who picked up my knife?’ (Valentine 2001:129)

² According to Valentine (2001:17), “Nishnaabemwin is a recent development of several distinct varieties of Ojibwe that were spoken on the shores of Lake Huron and Georgian Bay, and inland points to the east. Historically, speakers of Nishnaabemwin are descended from groups identified in English as Chippewas, Ottawas, Odawa, Potawatomi, and perhaps other indigenous peoples [...] In the most general terms, there exist two distinct sub-dialects within Nishnaabemwin, Odawa (Ottawa), spoken in Michigan and along the shores of Lake Huron, and Eastern Ojibwe, spoken in the area to the east, and including especially the communities of Rama and Curve Lake”.

³ The variation in the forms is dialectal. The final *-sh* comes from the fused contrastive focus particle *dash* (Valentine 2001:129, 979).

- (2) *wegnesh wezhtooyan?*
 IPW.INAN 2SG>3INAN.is.making
 ‘What are you making?’ (Valentine 2001:129)
- (3) *wenesh dnawa bneshiinh?*
 IPW.AN that.kind.AN.SG.PROX bird[AN.SG.PROX]
 ‘What kind of bird is it?’ (Valentine 2001:980) (? lit.: ‘What is that kind of bird?’)

Grammatically animate nouns “typically refer to living things, such as people, animals, spiritual beings, fish, birds, trees, and many plants”, whereas grammatically inanimate nouns “typically refer to non-living things, such as chairs, tables, boats, and books, but include certain plants as well” (Valentine 2001:114). However, “there is also a substantial group of nouns that refer to things that are not usually considered to be ‘living’, but which are grammatically animate”, e.g. *sab* ‘net’, *kik* ‘kettle, ashcan, pot’, *daabaan* ‘car’, words referring to heavenly bodies, processed lumber and wooden objects, grains and grain products, tobacco and related paraphernalia, hides, etc. (Valentine 2001:114-118). This should imply that while the inanimate interrogative pronominal *wegnesh* always means ‘what?’, the animate interrogative pronominal *wenesh* may mean both ‘who?’, when used in questions about humans, and ‘what?’, when used in questions about animals (and perhaps some plants), which are then ANIMATE-questions with ‘who?’-dominance (cf. Section II.4), and in questions about some grammatically animate non-living things, which would be an instance of lack of differentiation between ‘who?’ and ‘what?’. Thus, Valentine (2001:979) translates *wegnesh* as ‘what? (INAN)’ and *wenesh* as ‘who?/what? (AN)’. Moreover, according to Valentine (2001:129), “some speakers use *wegnesh* [IPW.INAN] with animates as well, and some use *wenesh* [IPW.AN] with inanimates. There is much variation”. Example (4) appears to provide a good illustration for the point at issue.

Nishnaabemwin

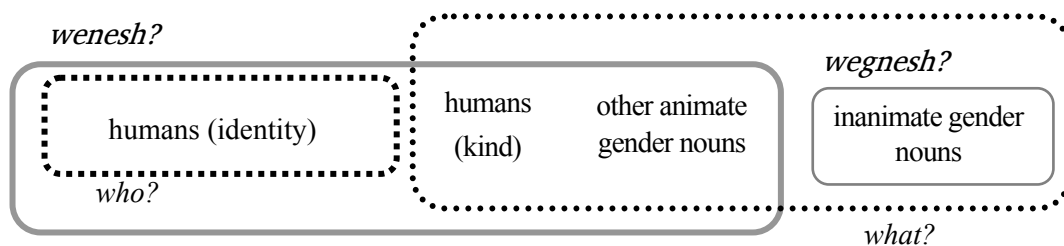
- (4) [*Giimooj sii go ngii-wiindmaag Linda baabiitood wii-ggwejmad...*]
wii-ggwejmag wenesh?
 1SG>3SG.AN.PROX.ask IPW.AN
 ‘[A: Linda told me in secret that she was waiting for you to ask her... B:]
 Ask her **what?**’ (Valentine 2001:991)

Interestingly, the animate vs. inanimate distinction appears to be blurred even more commonly in the free-choice indefinite pronominals, which seem to be related to the interrogative pronominals. Thus, “for many speakers”, the animate form *wegwen(h)/wegwendig* “has become the general form of the dubitative

pronoun [i.e., the free-choice indefinite pronoun], both animate and inanimate”, ‘whoever, whatever’ (Valentine 2001:127).⁴

I supposes the blurring of the animate vs. inanimate distinction in the interrogative pronominals (and free-choice indefinites) may be explained by an interference from English, viz. the mismatch between the patterns of use of the animate *wenesh* and human *who?*, on the one hand, and the inanimate *wegnesh* and the non-human ‘what?’, on the other, as schematized in Figure 1.

Figure 1. English *who?* and *what?* vs. Nishnaabemwin *wenesh* IPW.AN and *wegnesh* IPW.INAN⁵



Given that in Nishnaabemwin the animate interrogative pronominal *wenesh* corresponds both to English *who?* and, in certain contexts, to English *what?*, some Nishnaabemwin speakers might have reinterpreted the original animate interrogative *wenesh* as the equivalent not only to the English *who?*, but also *what?* and not only in questions about non-human grammatical animates (or KIND-questions), but also, apparently by extension, in regular questions about any non-humans. As to the Nishnaabemwin speakers that “use *wegnesh* [IPW.INAN] with animates as well”, the reinterpretation must have occurred, as it were, in the opposite direction. Thus, first, the inanimate interrogative *wegnesh* might have been interpreted as the equivalent to the English *what?*. However, since *what?* is used in English in several contexts where Nishnaabemwin would require the animate interrogative *wenesh*, the original inanimate interrogative *wegnesh* might have been reinterpreted as the animate interrogative pronominal as well.

Such presumably contact-induced change in Nishnaabemwin would be comparable to the similar development suggested in Section III.3.5.2 for a variety of the Tungusic language Evenki. In the case of the Evenki variety at

⁴ The original inanimate form of the free-choice indefinite is *wegdagwen/wegdogwen* ‘whatever’ (Valentine 2001:127).

⁵ The cell “humans (kind)” stands for KIND-questions (Section II.2) such as English *What is John?* (a lawyer, a doctor, etc.). My inclusion of this cell under the animate interrogative pronominal *wenesh* is based on indirect evidence only. Thus, in other Algonquian languages, such as Arapaho (Andrew Cowell, p.c.) and Cheyenne (Wayne Leman, p.c.), the inanimate interrogative cannot be used in questions involving a non-prototypical combination of values of the kind [person + classification (+ common noun)].

issue, the blurring of the human vs. non-human distinction in the interrogative pronominals may be attributable to interference from Russian.⁶ Although unlike Evenki, Nishnaabemwin is probably not an immediately endangered language, (almost) all speakers of Nishnaabemwin are bilingual in English and are exposed to everyday English influence.

By way of conclusion, a few words can be said about other Algonquian languages. As already mentioned, other Algonquian languages are similar to Nishnaabemwin in distinguishing two genders, animate and inanimate, and in making the same distinction in their interrogative pronominals. What is more, just like in Nishnaabemwin, the animate gender in other Algonquian languages typically contains a substantial group of nouns that “refer to things that are not usually considered to be ‘living’”. However, this does not seem to imply automatically that the animate interrogative can be used not only in questions about persons as ‘who?’ but also in questions about animals (and some plants), which would then be ANIMATE-questions with ‘who?’-dominance (cf. Section II.4), and in questions about some grammatically animate non-living things, which would be an instance of lack of differentiation between ‘who?’ and ‘what?’. Thus, in Arapaho (Andrew Cowell, p.c.) and Cheyenne (Wayne Leman, p.c.), the respective animate interrogatives appear to be restricted as ‘who?’ to questions about humans.⁷ At the same time, some Algonquian languages appear to be similar to Nishnaabemwin. For instance, Southwestern Ojibwa, as spoken in Ponemah, Minnesota, is reported by Black (1971:148) to have the following interrogative pronominals: *awegonen* glossed as ‘what inanimate gender thing (*What X...?*)’ and *awenen* glossed as ‘what animate gender thing (Usually, *Who...?*)’. However, as was mentioned in Section II.4.2.2, according to Schwartz & Dunnigan’s (1986:304) description of the pronominals of the same idiom, the animate interrogative *awenen* is rather restricted to questions about humans and “large animal[s], such as a bear or a horse”, as in (5a, b), while the inanimate interrogative, which they spell *awekonen*, would be used for insects for instance, as in (5c), even though e.g. *sakime* ‘mosquito’ is animate.

⁶ Interestingly, the original mismatch between the patterns of use of the interrogative pronominals of Evenki and Russian is very much the reverse of that between Nishnaabemwin and English. That is, Russian is more like Nishnaabemwin and Evenki is more like English.

⁷ For instance, Wayne Leman (p.c.) reports on Cheyenne that he has “never heard anyone question about an animal with *nevaahé*”, which is the animate interrogative pronominal. It is possible that the exclusively human meaning ‘who?’ of the animate interrogatives in languages like Cheyenne and Arapaho could be due to English influence as well, since in English *who?* is restricted to questions about humans.

Southwestern Ojibwa of Ponemah, Minnesota (Algic, Central Algonquian, Ojibwa; USA; Schwartz & Dunnigan 1986:304)

- (5) a. *awenen kaa-takkwamaat?*
 IPW.AN PST-bite.DIRECT.3OBJ
 ‘Who (person) bit him/them (person)?’
- b. *awenen-an kaa-takkwamikot?*
 IPW.AN-OBV PST-bite.INVERSE.3OBJ
 ‘What/who (“a large animal, such as a bear or a horse”, rather than a person) bit him/them (person)?’
- c. *awekonen kaa-takkwamikot?*
 IPW.INAN PST-bite.INVERSE.3OBJ
 ‘What (e.g., an insect) bit him/them (person)?’

Interestingly, with the animate interrogative, sometimes the human vs. non-human distinction can still be expressed indirectly, viz. “the inverse form [with the obviative marking on the animate interrogative, as in (5b)] is more appropriate for non-human A acting on human P, while the direct form [as in (5a)] is more appropriate if A and P are both human” (Schwartz & Dunnigan 1986:304).

Cyr (1992) describes Montagnais as having two interrogative pronominals, *auen* ‘who?, what? (animate)’ and *tshekuan* ‘who?, what?, why? (inanimate)’⁸ and states that the interrogative pronominal takes the gender and the number of the questioned entities (“prend le genre et le nombre des entités à propos desquelles on interroge”), as in (6). I have preserved the original French translation.

Montagnais (Algic, Central Algonquian, Cree; Canada; Cyr 1992)

- (6) a. *tshekuan ne?*
 IPW.INAN DEM
 ‘Qu’est-ce que c’est? (la réponse attendue est un objet inanimé)’ (‘What is this? (the expected answer is an *inanimate* object)’)⁹

⁸ Cyr’s use of the gloss ‘who?’ for the inanimate interrogative *tshekuan* is not completely clear to me. The inanimate gender would not normally be expected to include any human nouns (cf. Cyr 1992: “Si les noms d’êtres humains et d’animaux sont de genre animé, pour le reste on ne voit pas toujours ce qui motive l’appartenance d’un nom à un genre”). An interpretation that suggests itself is that just like in Nishnaabemwin, in Montagnais the animate/inanimate distinction in the interrogative pronominals is being blurred, presumably under the influence of French, which is quite similar to English in its patterns of use of the interrogative pronominals.

⁹ Cyr (1992) does not explain the use of the demonstrative *ne* here, which, according to her own description, is an *animate* singular demonstrative (cf. e.g. Cyr 1992, example (40)).

b. *auen ne?*

IPW.AN DEM

‘Qu’est-ce que c’est? (la réponse attendue est un objet animé)’ (‘What is this? (the expected answer is an *animate* object)’)

Similarly, in Maliseet-Passamaquoddy/Malecite-Passamaquoddy, it seems possible to use the animate interrogative pronominal *wen*¹⁰ in questions about humans (7), animals (8b), and (at least in predicative use) about grammatically animate things with non-live referents (9).

Maliseet-Passamaquoddy/Malecite-Passamaquoddy (Algonquian; Canada & USA)

(7) *wen-ik yukk ketuwahkatomuhtit?*

IPW.AN-AN.PL DEM.AN.PL who.want.to.get.married

‘Who are these people who want to get married?’ (Francis & Leavitt 2007, under *kotuwahkatom*)

(8) a. *Albert kete nehpah-a-l otuhk-ol*

PROP for.example kill-DIRECT-3OBV deer-OBV

‘Albert for example killed a deer?’ (Bruening 2001:77)

b. *Albert kete wen-il nehpah-ac-il*

PROP for.example IPW.AN-OBV kill-3PROX.CONJUNCT-PTCP.3OBV

‘Albert for example, what did he kill?’ (Bruening 2001:77)

(9) *wen-ik lahtu-wok, tolepsis-ok kosona kalus-iyik?*

IPW.AN-AN.PL trump-AN.PL club-AN.PL or diamond-AN.PL

‘What’s trump (lit.: ‘trumps’), clubs or diamonds?’ (Francis & Leavitt 2007, under *kalus*)

7.3 Timucua

Timucua used to be the primary native American language of southeastern Georgia and northern Florida. It became extinct somewhere in the late 1700s or early 1800s. All modern accounts of this language, such as Gatschet (1877, 1878, 1880) and Granberry (1990, 1993), are based almost exclusively on the early 17th century description, vocabulary and texts produced by Spanish missionaries (cf. Granberry 1990:61-62 or Mithun 1999:519-520 for more details and references). Timucua is believed to have comprised eleven distinct varieties. However, almost all extant data come from only two varieties, Mocama and Potano (Granberry 1990:61-62, 1993:7).

As far as its linguistic affiliation is concerned, for the moment Timucua appears to be best treated as an isolate. However, Julian Granberry (1993:15, 59)

¹⁰ Its inanimate counterpart is *keq/keqoss/keqsey* ‘what?’ (cf. Francis & Leavitt 2007).

argues that Timucua is in origin a “creolized” idiom, largely based on some South American source. Thus, he suggests Warao, an isolate spoken in the Orinoco Delta in Venezuela and neighbouring regions of Guyana, as the closest relative, although he also claims that there is “an even larger number of lexemes with equally striking resemblance to languages of the Vaupés-Caquetá-Inírida-Guaviare branch of Northern Maipuran Arawakan”. In Granberry’s view, the “ultimate origin for the [Timucua] language and people [must be] somewhere in the Vaupés-Caquetá region of Colombia”. Given that the North Arawakan languages were spoken throughout the Caribbean in pre-Columbian times, where they had migrated from continental South America, Granberry’s hypothesis cannot be dismissed *a priori*.

The Timucua descriptions at my disposition, Gatschet (1877, 1878, 1880) and Granberry (1990, 1993), treat interrogative pronominals only marginally. Gatschet (1877, 1878, 1880) is much better as far as sentential examples are concerned. Granberry (1990, 1993) provides hardly any sentential examples, but Granberry (1993) contains a Timucua-English wordlist and an English-Timucua finder list. Table 1 reproduces the interrogatives glossed as ‘who?’, ‘what?’, ‘which?’ and ‘where?’ in the sources. Note that the wordlist in Granberry (1993) has a lot of forms translated as ‘who’, ‘what’, ‘which’ or ‘where’. However, here I included only the forms for which the glosses involved a question mark, i.e. ‘what?’ and not ‘what’, or for which the interrogative function was explicitly mentioned.

The final *-co* in some of the forms is one of the following: (i) *-co* “interrogative post-clitic” (other forms being *-che* and *-cho*), (ii) *-co* “generalizing post-clitic – ‘the one who...’” (i.e. a nominalizer), (iii) *-co* “‘may, might’ (= conditional action)”, (iv) *-có(co)* “1. copular post-clitic; 2. *By extension* an augmentative/intensive post-clitic”, (v) *co* “to say” (cf. Granberry 1993:122-125). The variation in the quality of the morpheme-final vowels, as in Gatschet’s *hacha-/hachi-*, is an extremely common process in Timucua and is mostly morphonologically conditioned (cf. Granberry 1990, 1993:68-76).

As can be observed in Table 1, Granberry gives two interrogative forms glossed as both ‘who?’ and ‘what?’, *cho-n-co* and *muchu*, while Gatschet gives only one, but with three variants, *hacha*, *acha* and *cha*. The obvious formal and semantic disagreement between the two sources cannot be solved with the data available, because Gatschet (1877, 1878, 1880) provides only examples for the non-human uses of *((h)a)cha* and Granberry (1993) provides no sentential examples with interrogatives. Furthermore, given the *hacha-/hachi-* variation and given that in *cho-n-co* the second syllable is originally *-no*, it is tempting to hypothesize that both the human interrogative root *chi-* ‘who?’ and the root *cho-* of *cho-n-co* are just morphonological variants of the same interrogative root *((h)a)cha*. Should all these forms be indeed related, from the semantic point of

Table 1. Timucua interrogatives ‘what?’, ‘who?’, ‘which?’ and ‘where?’ (based on Gatschet 1877, 1878, 1880; Granberry 1990, 1993)

Granberry (1993)	Gatschet (1877, 1878, 1880)
<i>hacha</i> “general interrogative = what, why, how, where, whence, whither”, also ‘something, someone; anything, anyone; property; faithful’ (p.132)	<i>hacha, acha, cha</i> ‘who?, what?, which?’ (1877:632)
<i>cha</i> ‘what?’; ‘where, whence, whither’ (p.122)	
<i>cho-n-co</i> (< <i>cho-no-co</i>) ‘what?, who?’ “(lit. ‘who may it be?’)” (p.124)	<i>cha-n-co</i> (< <i>cha-na-co</i>) ‘whence?, where from?’ (1877:632)
<i>hachibono</i> ‘that; thing’	<i>hachi-/hacha-</i> + <i>-mueno/</i> <i>-mono/-bueno/-</i> <i>buena/-bono</i> ‘what?, which thing?, which?, something’ (1877:631, 641, 1888:497-501) ‘name’
<i>michu</i> ‘what?, which?, who?’, also ‘that (one)’ (p.148)	
<i>chita-co</i> ‘who?’ (p.124)	<i>chi-ta, chi-ta-co-</i> ‘who?’ (1877:634, <i>ca-re-n-te</i> (‘who 1888:502, 1880:494) are [they]?’), <i>chi-</i> <i>ca-co-n-te</i> (‘who is [3SG]?’)

view the most plausible original meaning would have been the locative ‘(be) where?’, with the pronominal meanings having evolved through the selective use of ‘(be) where?’ as ‘which one? (person or thing)’.

7.4 Timbisha

The Timbisha language, also known as Panamint and Timbisha/Tümpisa Shoshone, is a moribund language of the Central Numic subgroup of the Northern Uto-Aztecan languages. Traditionally, Timbisha was spoken in what is today southeastern California and southwestern Nevada. The Timbisha territory included Death Valley and the surrounding valleys, deserts and mountain ranges. According to Dayley (1989:6), “in earlier times, people from different valleys,

even people from different villages within the same valley, spoke somewhat different dialects”. McLaughlin (2006:3) speaks of three major dialects: (i) Eastern Timbisha, spoken by the remnants of the Grapevine Canyon and Beatty Timbisha communities, and described by McLaughlin (2006), (ii) Central Timbisha, represented by the Death Valley variety of Timbisha, and described by Dayley (1989), (iii) Western Timbisha, covering the dialects from the Coso Region and Owens Valley. Given the lack of data, I will disregard Western Timbisha in what follows. Of the two remaining varieties, only Eastern Timbisha appears to have a full-fledged ‘who?, what?’ interrogative. Therefore, I will begin this section by presenting the Eastern Timbisha data. Then, the interrogative pronominals of Central Timbisha will be presented. Finally, I will briefly discuss the possible origins of the lack of differentiation between ‘who?’ and ‘what?’ in Eastern Timbisha.

Eastern Timbisha interrogative pronominals are reproduced in Table 2. Their use is illustrated in (10-13).

Table 2. Case-number paradigm of the Eastern Timbisha interrogative pronominals (based on McLaughlin 2006:25)

	NOM	ACC	POSS
SG	<i>hakatin</i>	<i>hakka</i> ‘who?, what?’ <i>hinna</i> ‘what?’ <i>hakatina</i> ‘which one? (person or thing)’	<i>hakkan</i>
DU	<i>hakatiŋku</i>	<i>hakatiŋkʷi</i>	<i>hakatiŋkʷin</i>
PL	<i>hakatiŋmi</i>	<i>hakatiŋmi</i>	<i>hakatiŋmin</i>

Eastern Timbisha (Northern Uto-Aztecan, Central Numic; McLaughlin 2006:25-26)

- (10) *hakatin* *u* *pakka-htai-mma?*
IPW.SG.NOM 3SG.ACC kill.one-COMPL-INDF.PST
‘Who killed him?’
- (11) *hakka* *i* *yaa-ttai-mma?*
IPW.SG.ACC 2SG.NOM take.one-COMPL-INDF.PST
‘What did you take?’
- (12) *hakatina* *i* *yaa-ttai-mma?*
which.one.SG.ACC 2SG.NOM take.one-COMPL-INDF.PST
‘Which one did you take?’
- (13) *hinna* *ni* *tikka-tu’ih?*
what.SG.ACC 1SG.NOM eat-FUT
‘What am I going to eat?’

Except *hinna*, all interrogative pronominals in Table 2 “are based on the stem *haka*” (McLaughlin 2006:25). “Interrogatives of location and direction are [also] formed by using the appropriate postposition on the interrogative base *haka*” (McLaughlin 2006:26), as in (14).¹¹

Eastern Timbisha (McLaughlin 2006:26)

- (14) *haka-pan ni kati-tu'ih?*
 IPW-on 1SG.NOM sit-FUT
 ‘What will I sit on?’

Unfortunately, McLaughlin (2006) does not mention whether it is also possible to use non-locative “postpositions” with *haka*-, such as the accompaniment postposition *-ma'ai* ‘with’ or instrumental *-man* ‘with’. This question is important, because for instance in Central Timbisha, this is possible and *haka* then may have a human meaning ‘who?’, as in (15a) vs. (15b).

Central Timbisha

- (15) a. *haka ma'e su-tü?*
 IPW with that.not.visible-NOM
 ‘Who was he with?’ (Dayley 1989:149)
 b. *haka-pai pittuhuntü?*
 IPW-from return
 ‘Where’s he returning from?’ (Dayley 1989:151)

It is also possible in Central Timbisha, but apparently not in Eastern Timbisha, to use the interrogative *haka* on its own, i.e. not marked with a case ending or with a postposition, either as a locative interrogative ‘(be) where?’ (16) or in questions about appearances (17a), things said or thought (17b), in which case it corresponds to English ‘how?’ or ‘what?’ (not about entities). However, *haka*-based derivatives, such as *haka-pa* ‘where?’, *haka-ni* ‘how?, in what way/manner?, (do) what/how?’, *haka-mi* ‘(say, think) what?/how?’, appear to be much more common in Central Timbisha.

Central Timbisha

- (16) *haka su-tü?*
 IPW that.not.visible-NOM
 ‘Where is it?’ (Dayley 1989:149)

¹¹ Most Eastern Timbisha “postpositions” are in fact affixes (cf. McLaughlin 2006:30-33: “the postposition is *affixed* directly to the noun”, italics added). However, in Dayley’s (1989) description of Central Timbisha the postpositions are more often written as words, not as affixes. For convenience sake, I will follow the sources in using the label postposition.

- (17) a. *haka napunni sa-tü?*
 IPW look that.visible-NOM
 ‘What does that look like?’ or ‘How does that look?’ (Dayley 1989:149)
- b. *haka usun takasü ni-a nasuntamana*
 INDF the.one.not.visible.that all 1SG-ACC remember-SUB
 ‘That’s all of what I remember’ (Dayley 1989:153)¹²

Table 3 summarizes the Central Timbisha interrogative pronominals “that have been recorded” (Dayley 1989:148). Therefore, the lack of dual and plural forms in Table 3 is probably accidental. I am not sure, though, whether the same accounts for the lack of the possessive form of ‘what?’.

Table 3. Central Timbisha interrogative pronominals (based on Dayley 1989:148)

	NOM	ACC	POSS	+ a postposition
‘who?’	<i>hakatün</i>	<i>hakka</i>	<i>hakkan</i>	<i>haka-</i>
‘what?’	<i>hii</i>	<i>hinna</i>	—	

As can be observed in Table 3, in Central Timbisha the lack of differentiation between ‘who?’ and ‘what?’ is only possible in non-core cases. It does not seem to be obligatory, though, given that Central Timbisha (similarly to Eastern Timbisha) can also build postpositional phrases using the accusative and nominative forms of nominals.

It is possible both in Eastern and Central Timbisha for the same forms to be used as both interrogatives and indefinites, although Central Timbisha would “more typically” use some extra morphology to mark the indefinite function more explicitly (cf. Dayley 1989:152-153; McLaughlin 2006:26-27). This would normally imply that when a given interrogative pronominal can have both a human and a non-human meaning, it should also be possible to use the same form indefinitely in both meanings. Unfortunately, the relevant examples are lacking in the sources.

By way of conclusion, let us consider the possible origins of the lack of differentiation between ‘who?’ and ‘what?’ in Eastern Timbisha. I believe that this lack of differentiation must be due to the fact that the original meaning of the stem *haka-*, on which almost all of its interrogative pronominals are based, was locative ‘(be) where?’, indifferent to the human vs. non-human distinction.

¹² Admittedly, this example is not a question. However, given that in Timbisha interrogatives and indefinites are expressed with the same forms, I have considered it possible to use this example here instead of the respective interrogative one, which unfortunately is lacking.

Subsequently, this locative use was extended to the selective ‘which one? (person or thing)’ and finally expanded to non-selective contexts as ‘who?, what?’, almost completely ousting the earlier stem for ‘what?’ **hi(k)i* (or the like). The latter has survived only as an option in the accusative, as well as in some derived non-pronominal interrogatives, such as *hinni* ‘(do) how?’ and *himpenni* ‘when?’ (McLaughlin 2006:26). In Central Timbisha, as in the other Numic languages, the case marked forms of the stem *haka* have been restricted to the human meaning ‘who?’ when expanding from the selective to the non-selective pronominal use, which is not surprising either, given the usual correlation between higher specificity and humanness. The same did not happen to the *haka* forms marked by a postposition because, I suppose, they have been perceived as circumstantial, adverbial, and therefore, comparable to the similar ‘where?’ and ‘how?’ uses of the unmarked *haka*. Note in this respect that the manner interrogatives meaning ‘how?’ in Numic languages seem to be typically based on *haka* rather than **hi(k)i*. Thus, the closely related Central Numic language Western Shoshoni appears to have only *haka-ni* ‘how?’ (cf. <http://www.shoshonidictionary.com>). Similarly, the Southern Numic language Ute has only *’agá-ni* ‘(do) how?’ and *’agá-rā’a* ‘(be) how?’ (cf. Givón 1984:230).¹³

Besides being semantically plausible, the hypothesis proposed above is further supported by the following facts. First, in Central Timbisha *haka* used on its own as a predicate means ‘be where?’.¹⁴ Second, in Eastern Timbisha, *haka-tín-a*, the regular accusative form of the nominative singular general interrogative pronominal *haka-tín*, has the selective meaning ‘which one? (person or thing)’. Finally, recall that in Numic languages the manner interrogatives meaning ‘how?’ appear to be typically based on *haka* (or the like), which squares well with the presumed original locative meaning of the latter.¹⁵

¹³ It should be mentioned that Ute *’agá-*, the “referential inanimate WH-pronoun”, seems to correspond to the Timbisha *hakka*, while Timbisha *haka* rather corresponds to the Ute nominative “referential animate WH-pronoun” *’aa* (for the forms see Givón 1984:226).

¹⁴ Its manner use as ‘how?’ is clearly secondary. The extension from ‘where?’ to ‘how?’ is not uncommon cross-linguistically. For instance, it can be found in quite a few Australian languages.

¹⁵ Interestingly, the locative origin of *haka* (or the like) can help to explain one apparent anomaly signalled by Givón (1984:228) for Ute. Thus, according to Givón, Ute uses *’agá-*, the “referential inanimate WH-pronoun” to “elicit adjectival predicate responses”, e.g. *What is this man like?*, *What is this land like?*, which is “not synchronically explicable”, “since these are obviously non-referential (‘type’) questions”. However, this synchronic anomaly is readily explainable if we assume that Ute *’agá-* originates in a locative interrogative ‘(be) where?’, which could also be used as ‘how?’. The rather natural semantic extension of ‘(be) where?’ to ‘which one?’, already discussed above, would also explain why Ute inanimate *’agá-* and its

7.5 Mayan languages¹⁶

Mayan languages are spoken primarily in Guatemala, Mexico and Belize. Depending on one's definition of language, the number of Mayan languages may range from around thirty (e.g., Campbell & Kaufman 1985) to around seventy (e.g., the Ethnologue). Here, I will generally adopt the more moderate classification. Sometimes, the Ethnologue's labels will be provided as well. According to the received view, the Mayan family comprises five branches, Huastecan, Yucatecan, Cholan-Tzeltalan/Greater Tzeltalan, Kanjobalan-Chujean/Greater Kanjobalan and Quichean-Mamean/Eastern Mayan, as illustrated in Figure 2. It is also commonly believed that the Huastecan branch "was the first to break off" the Proto Mayan (Campbell & Kaufman 1985:188), followed by Yucatecan. The Cholan-Tzeltalan and Kanjobalan-Chujean branches are sometimes grouped together into the Western Mayan branch, as opposed to the Quichean-Mamean/Eastern Mayan branch.

There appear to be at least four to five Mayan languages that allow for the lack of distinction between 'who?' and 'what?': Tz'utujil¹⁷ (some/?all dialects; Section III.7.5.1), Kaqchikel (most dialects; Section III.7.5.2), Mam (one dialect, but the data are inconclusive; Section III.7.5.3), Uspanteka (at least one dialect; Section III.7.5.4), and Tzotzil (at least one dialect; Section III.7.5.5). Tz'utujil, Kaqchikel and Uspanteka are Quichean, Mam is Mamean, and Tzotzil is a Tzeltalan language. In what follows, I will first present the interrogative pronominals of each language in more detail. By way of conclusion, a brief summary will be provided in Section III.7.5.6.

7.5.1 *Tz'utujil*

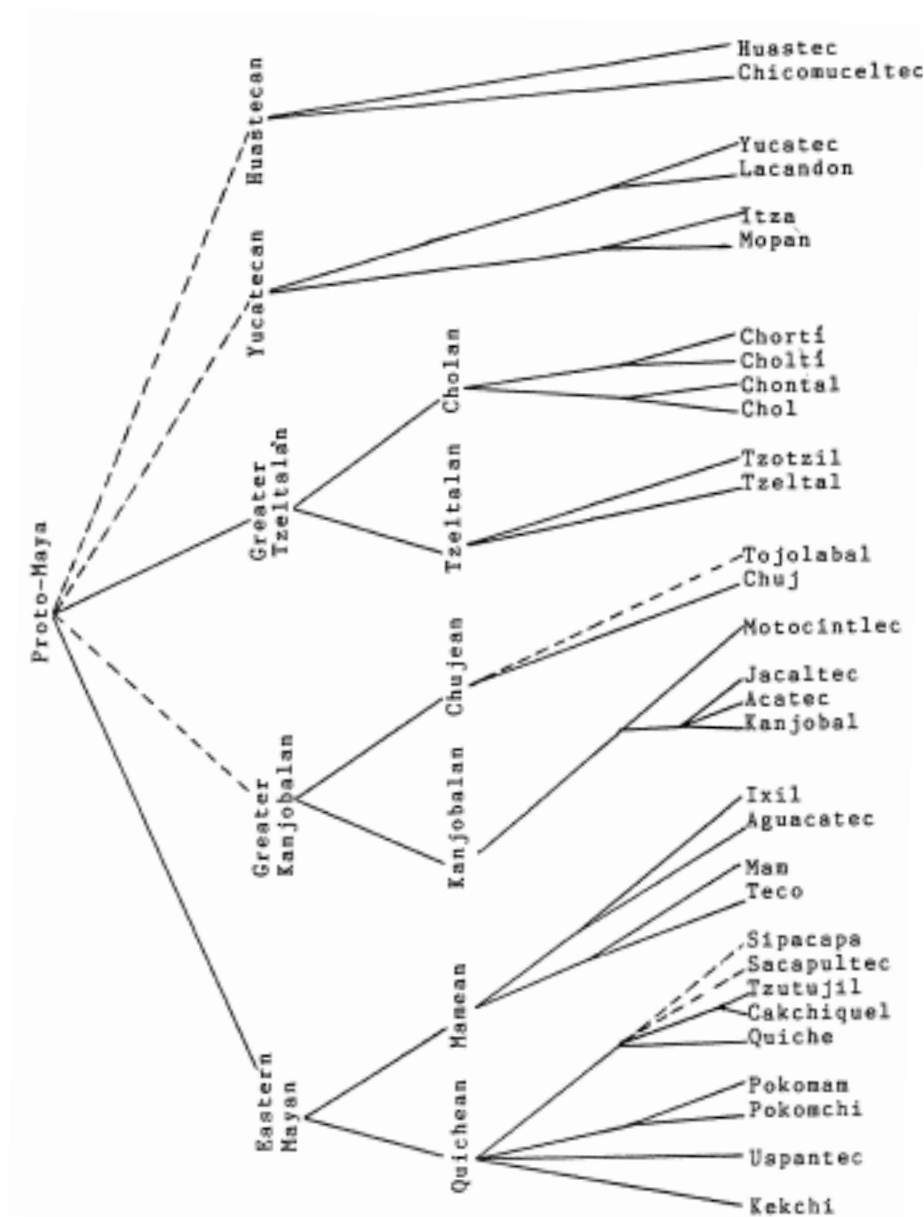
Tz'utujil/Tzutujil is a relatively small Quichean Mayan language spoken in southern Guatemala to the south of Lake Atitlán. According to Dayley (1985:1) there are eight Tz'utujil towns, as well as numerous villages. Tz'utujil is similar to many other Mayan languages in that "a different variety of Tzutujil is spoken

animate counterpart *ʔáa* have developed into "referential" interrogative pronominals (the "referential" form is used when "the speaker expects [a] unique referential identification, most likely a definite description", Givón 1984:226).

¹⁶ Whenever possible the spelling of examples from Mayan languages has been uniformized, primarily in accordance with the practical Spanish-based orthographies used for the Guatemalan Mayan languages. Thus, *k* stands for /k/, *q* for /q/, *ch* for /tʃ/, *x* for /ʃ/, *tz* for /ts/, *j* is normally /x/, the glottal stop and glottalization of consonants are represented with an apostrophe, as in *a'* or *q'*.

¹⁷ The names of the Guatemalan Mayan languages are spelled following the Guatemalan Academy of Mayan Languages (see <http://www.almg.org.gt>).

Figure 2. Classification of Mayan languages (Campbell & Kaufman 1985:189)¹⁸



in virtually every town in the area [...] although none of [the] differences are so great that any of the varieties are mutually unintelligible” (Dayley 1985:3). The Ethnologue subdivides Tz’utujil into two languages Western Tz’utujil and Eastern Tz’utujil. The interrogative pronominals used both as ‘who?’ and ‘what?’ are reported for both Western and Eastern Tz’utujil. Thus, Kaufman (1976 via Dienhart 1997) reports the interrogative pronominal *naq* ‘who?, what?’ for the Santiago Atitlán dialect of (Eastern) Tz’utujil. Dayley (1985) provides a detailed

¹⁸ “Dotted lines represent less secure or more controversial groupings” (Campbell & Kaufman 1985:189).

description of a similar ‘who?, what?’ interrogative in the San Juan la Laguna dialect of (Western) Tz’utujil. I do not have information on the other dialects.

Let us consider the interrogative pronominals of San Juan la Laguna Tz’utujil in more detail. According to Dayley (1985:69-70, 238-240, 331-335), San Juan la Laguna Tz’utujil has two interrogative pronominals *naq* ‘who?, what?, what/which [N]?’ (18-21, 23) and (*naq*) *choq*/*(naq)* *choj* ‘whom?, what?’ (21-22).

San Juan la Laguna Tz’utujil (Mayan, Quichean; Guatemala)

- (18) a. *naq aa-wach?*
 IPW 2SG.GEN-face
 ‘Who are you?’ (lit.: ‘Who/what is your face/surface/character/appearance/being/type/kind?’) (Dayley 1985:237, 239)
- b. *naq r-wach?*
 IPW 3SG.GEN-face
 ‘Who/what is it?’ (Dayley 1985:237)
- c. *naq awa’?*
 IPW DEM
 ‘What is this?’ (Dayley 1985:238)
- (19) a. *naq x-Ø-a-b’an-b’e-ej?*
 IPW COMPL-3SG.OBJ-2SG.A-do-APPL-NON<PRF>
 ‘What did you do it with?’ (Dayley 1985:237)
- b. *naq x-aa-tz’at?*
 IPW COMPL-3SG.OBJ-2SG.A-see
 ‘Who/what did you see?’ (Dayley 1985:237)
- (20) a. *naq n-Ø-pit chwaaq?*
 IPW INCOMPL-3SG.S-come tomorrow
 ‘Who/what is coming tomorrow?’ (Dayley 1985:332)
- b. *naq x-Ø-tij-ow-i*
 IPW INCOMPL-3SG.OBJ-eat-ANTIPASSIVE-NON<PRF>
ja w-ajaache’l?
 the 1SG.GEN-white.zapote
 ‘Who/what ate my white zapote?’ (Dayley 1985:332)
- (21) a. *naq chee x-Ø-b’e?*
 IPW to COMPL-3SG.S-go
 ‘Why did she go?’ (Dayley 1985:335)
- b. *choq chee x-Ø-aa-ja’ wi’?*
 IPW to COMPL-3SG.OBJ-2SG.A-give OBL.FOC
 ‘To whom did you give it?’ (Dayley 1985:240)

- c. *choq chee x-Ø-aa-choy wi'?*
 IPW with COMPL-3SG.OBJ-2SG.A-cut OBL.FOC
 'With what did you cut it?' (Dayley 1985:240)
- (22) *(naq) choq xiin awa' sipp ri'?*
 IPW IPW of/for DEM present DEM
 'For whom/whose is this present here?' (Dayley 1985:333)
- (23) a. *naq jaay k'aari'?*
 IPW house that
 'Which house is that?' (Dayley 1985:332)
- b. *naq óora x-ee-b'e?*
 IPW hour COMPL-3PL.S-go
 'What time did they go?' (Dayley 1985:334)

The first interrogative, *naq*, is used "to question direct arguments in a proposition, that is, subjects of intransitive verbs and stative predicates, and agents and patients of transitive verbs", while *(naq) choq* is used "to question the following oblique arguments: datives, instrumentals, benefactives, comitatives, and possessors" (Dayley 1985:69). The interrogative *(naq) choq* is "always used in conjunction with a following relational noun [as *chee* in (21) or *xiin* in (22)], which distinguishes the semantic role of the oblique argument" (Dayley 1985:69).¹⁹ However, the distribution between *naq* and *(naq) choq* appears to be somewhat more complex. Thus, *naq* can be used with some relational nouns as well, as in (21a). A few other combinations involving *naq* that may be worth mentioning are summarized in (24).

San Juan la Laguna Tz'utujil (Dayley 1985:151, 239)

- (24) a. [*naq* + the plural word *taq*] 'who all?, what all?'
 b. [*naq* + *chi* 'at, to, with' + abstract noun N] 'what kind of [N]?', as in
naq chi winaqiil 'what kind of person (is s/he)?' (*winaq* 'person, people' > *winaq-iil* 'humanness, naturalness')
 c. [*naq* + *chi-k-e* at/to/with-3PL.POSS-RN 'to them' + N] 'which one(s) is(are) N?'

¹⁹ The so-called "relational nouns" in Mayan languages are a group of nominals that function mostly as adpositions or conjunctions. They are often related to nouns designating body parts and are typically marked by a possessive affix for the person-number of their "complement", as in Tz'utujil *r-umaal jar aachi* (3SG.POSS-by the man) 'by the man', lit. 'his-by the man', where *-umaal* is the relational noun 'by', also meaning 'cause, fault' (Dayley 1985:152). When the possessive affix is lacking (or frozen), one usually speaks about adpositions (and conjunctions). However, for convenience sake, I will use the term relational noun throughout.

The possibility of *naq choq* next to simply *choq*, as in (22), seems to suggest that *choq* itself is in origin some kind of relational noun, that has become frozen in combination with the interrogative *naq*, with the latter being subsequently clipped off. More specifically, *choq* is probably a fused combination of *chi* ‘at, to, with’ + *q-e* 1PL.GEN-RN, similar to *chi-k-e* in (24c). Note, in this respect, that, first, *chi* ‘at, to’ is “sometimes [...] realized as *cha* before *q*”, as in *cha-q-e* ‘to us’ (Dayley 1985:155-156). Compare also the “instrumental” ‘(with) what?, (by) whom?’ and “dative” ‘[give] (to) whom?’ interrogative in the rather closely related Kaqchikel: *achoj* in most dialects vs. *chaq* in San Lucas Tolimán dialect (Pérez *et al.* 2000:125; Section III.7.5.2). Second, the default use of the first person plural with relational nouns (in conjunction with an interrogative) is attested in other Quichean-Mamean Mayan languages. Thus, in southern and central varieties of Mam, such as Cajolá, the meaning ‘by whom?’ (agent) or ‘with what?’ (instrument) is expressed by the combination [*al* IPW + *q-u’n* 1PL.GEN-RN] (cf. Pérez *et al.* 2000:108, 263-264).

7.5.2 Kaqchikel

Kaqchikel/Cakchiquel is a big Quichean Mayan language, rather closely related to Tz’utujil and spoken immediately to the north and north-east of the latter. The Ethnologue distinguishes ten Kaqchikel languages: Central Kaqchikel, Southcentral Kaqchikel, Eastern Kaqchikel, Northern Kaqchikel, Western Kaqchikel, Southern Kaqchikel, Akatenango Southwestern Kaqchikel, Santa María de Jesús Kaqchikel, Santo Domingo Xenacoj Kaqchikel, and Yepocapa Southwestern Kaqchikel. However, just like in the case of Tz’utujil, in almost every town a somewhat different variety of Kaqchikel appears to be spoken. Patal Majzul *et al.* (2000) offer a detailed comparative survey of a representative set of the Kaqchikel idioms. This survey also examines the variation in the interrogative pronominals in Kaqchikel. The results are summarized in Table 4.²⁰

²⁰ It should be pointed out that that in Patal Majzul *et al.* (2000), there exist certain inconsistency between the explicit textual presentation on pp. 73-76, 125-128 and the Swadesh lists for the respective dialects on pp. 180-189. The inconsistencies are both formal and semantic. Thus, the Swadesh lists show a more extensive formal variation and more dialects with distinct ‘who?’ and ‘what?’ than explicitly suggested in the text. First of all, this may be attributable to the way data are collected for such a list: (i) the words are translated from Spanish, (ii) the entries for ‘who?’ and for ‘what?’ are situated very close to one another, viz. ‘who?’ is the first entry and ‘what?’ is the fifth one. Furthermore, at least as far as the interrogative *achike* (and the like) is concerned, “there is a lot of internal variation in every community” (Patal Majzul *et al.* 2000:73). When the text and the data in the respective Swadesh list are inconsistent, I have adopted the textual explanation of the meanings of the interrogatives, but I provide all the forms mentioned.

Table 4. Interrogative pronominals in Kaqchikel dialects (based on Patal Majzul et al. 2000:73-76, 125-128, 180-189)

	A, S, OBJ			other functions
	‘who?, what?’ (also ‘which/what [N]?’)	‘what?’	‘who?’	
Patzicía	<i>(a)chike, chike’</i>			} <i>achoj</i>
Patzún	<i>ach(i)ke</i>	<i>atux</i> (elders only)	<i>aku’ux</i> (mostly PL and elders only)	
San Antonio Aguas Calientes	<i>ach(i)ke, anchke</i>		<i>an(ch)ku’x</i> (mostly PL)	
San Antonio Palopó	<i>achka, (a)ch(i)ke</i>	<i>atux</i>		
San José Poauquil	<i>chike</i>			
San Juan Sacatepéquez, ²¹	<i>(a)chike</i>			
San Miguel Pochuta				
San Marcos la Laguna	<i>naq (chi)</i>	<i>naxwäch, naqwäch</i>	<i>najwäch</i>	
San Martín Jilotepeque	<i>achike</i>	<i>axtu’atux</i> (elders only)		
San Pedro Sacatepéquez	<i>(a)chike, chke</i>			
San Pedro Yepocapa	<i>anchike</i>	<i>anux</i>		
Santa Catarina Palopó	<i>achike</i>			
Santa María Cauqué	<i>(a)chike, achinaq</i>			
Santa María de Jesús	<i>chika</i>			
Santo Domingo Xencoj	<i>achike(’), chike</i>			
Sololá	<i>achika</i>	<i>atux</i> (for some speakers)	<i>aku’(u)x</i> (mostly PL)	
Tecpán	<i>ach(i)ke, chke</i>	<i>atux</i> (elders only)	<i>aku’ux</i> (mostly PL and elders only)	
San Lucas Tolimán	<i>(a)chike, chke</i>			<i>chaq</i>
San Pedro Chuarrancho	<i>achike(’)</i>			<i>ache</i> “instrumental”, <i>apex</i> “dative”
San Andrés Semetabaj	—	<i>atux</i>	<i>achike</i> (SG) <i>aku’(u)x</i> (PL)	<i>achoj</i>

²¹ For San Juan Sacatepéquez, Stoll (1884 via Dienhart 1997) gives *qaxchiké* ‘who?’ and *qax* ‘what?’ (cf. Section III.7.5.6).

The dialects have mostly been ordered alphabetically. The column “other functions” covers those syntactic contexts where the interrogative pronominal needs to be marked with a postposed relational noun, e.g. “instrumental” ‘(with) what?, (by) whom?’ and “dative” ‘[give] (to) whom?’, etc.

As can be observed in Table 4, San Andrés Semetabaj Kaqchikel appears to be the only dialect that obligatorily distinguishes between human and non-human interrogative pronominals, at least in the core syntactic functions (A, S, OBJ), since when the interrogative pronominal needs to be marked with a postposed relational noun a single form *achoj* ‘who?, what?’ seems to be used. Moreover, there are several dialects that have a dedicated interrogative ‘what?’ (usually *atux*)²², next to a ‘who?, what?’ interrogative, and some dialects that also have a dedicated interrogative ‘who?’ (or rather ‘who? (PL)’), usually *aku(u)x*. The San Marcos la Laguna Kaqchikel forms *naxwäch/naqwäch* ‘what?’ and *najwäch* ‘who?’ look very much like contractions of an interrogative sentence ‘who/what is it?’, similar to (18b) in San Juan la Laguna Tz’utujil. Kaqchikel *wäch* appears to correspond to the Tz’utujil *wach* ‘face, surface, character, appearance, being, type, kind’.

The interrogatives *achoj* and *chaq* must have the same origin as the respective Tz’utujil interrogative *choq/choj* (cf. Section III.7.5.1). In its origins, the interrogative *ache* seems to differ from *achoj* and *chaq* only by the original use of the 3SG.GEN affix instead of the 1PL.GEN. In this respect, compare the Tz’utujil form *chee* ‘to/at/with him/her/it’ < *[*chi* ‘at, to, with’ + *r-e* 3SG.GEN-RN] (cf. Dayley 1985:156). The *-pex* part of the San Pedro Chuarrancho “dative” interrogative *apex* clearly contains *pa* ‘in, to’. The origin of the *-(e)x* part is not immediately clear.

The Kaqchikel ‘who?, what?’ interrogative *achike* and the like appears to represent a contraction of the earlier construction **a(n)chinaq chike* (or the like; cf. also Sections III.7.5.3 and III.7.5.6) ‘which one?’, which can be compared to the San Juan la Laguna Tz’utujil interrogative construction [*naq* IPW + *chi-k-e* at/to/with-3PL.POSS-RN ‘to them’ + N] ‘which one(s) is(are) N?’ (24c). This hypothesis is supported by the following facts (i) the existence of ‘who?, what?’ forms such as San Marcos la Laguna Kaqchikel *naq*, San Pedro Yepocapa Kaqchikel *anchike* and Santa María Cauqué Kaqchikel *achinaq* next to *(a)chike*; (ii) the existence of similar contractions in Kaqchikel with other interrogatives, as with *achike (r-)uma* ‘why? (lit.: IPW (3SG.GEN-)cause)’ that can also be

²² In fact, there may be more dialects having *atux* (or the like) as ‘what?’, because according to Patal Majzul *et al.* (2000:126), besides San Antonio Palopó, Sololá and San Andrés Semetabaj, “*atux* is also used in the other communities [“en las otras comunidades”] but only by elderly people”. However, elsewhere, Patal Majzul *et al.* (2000:74) speak only about two communities (not all “the other communities”), viz. Patzún and Tecpán, where *atux* ‘what?’ is still in use only by elderly people.

contracted to *aruma* ‘why?’ (García Matzar & Rodríguez Guaján 1997:212); (iii) the original *selective* meaning ‘which one? (person or thing)’ of **a(n)chinaq chike* (lit.: ‘who/what/which.one at/to them’) would square well with the fact that in San Andrés Semetabaj Kaqchikel *achike* has become specialized as a human interrogative ‘who?’.²³ Given that Tz’utujil is the closest relative of Kaqchikel, the Tz’utujil general interrogative pronominal *naq* ‘who?, what?, which/what [N]?’ must also go back to something like **a(n)chinaq* ‘who?, what?, which/what [N]’, which should then be reconstructed also for the common ancestor of Kaqchikel and Tz’utujil.

7.5.3 Mam

Mam is a big Mamean Mayan language, spoken primarily in southwestern Guatemala. The Ethnologue distinguishes five Mam languages: Central Mam, Northern Mam, Southern Mam, Tajumulco Mam, and Todos Santos Cuchumatán Mam. There is a great degree of dialectal variation within Mam. A comparative survey of a large part of the Mam idioms can be found in Pérez *et al.* (2000). It should be mentioned, though, that the northern varieties spoken in the western Huehuetenango and northern San Marcos Departments are best represented in Pérez *et al.*’s (2000) sample. This survey also examines the variation in the interrogative pronominals in Mam. In Pérez *et al.*’s (2000) sample, there is only one (northern) Mam idiom, spoken in the San Pedro Necta municipality, that uses a single interrogative *alche(e)* ‘who?, what?’ in A, S and OBJ functions (2000:34, 109-110). However, it is not clear from the description in Pérez *et al.* (2000) whether the same lack of differentiation between the human and non-human meanings of the interrogative pronominal is also possible in other syntactic functions, viz. when the interrogative needs to be marked by a postposed relational noun. For instance, Pérez *et al.* (2000:109) give the San Pedro Necta form *te qu’n* translated only as ‘by whom?’, which in all probability

²³ The final glottal stop in forms such as Patzicía Kaqchikel *chike* ‘who?, what?’ is probably a trace of some fused demonstrative. In Kaqchikel, normally the proximate demonstrative has the form *re*’ (and the like) and the distant demonstrative the form *la*’ (and the like), i.e. both with a final glottal stop. Note also that, for instance in the Swadesh lists in Patal Majzul *et al.* (2000:180-189), the San Pedro Yepocapa Kaqchikel interrogative ‘what?’ was given as *anux re*’, literally ‘what (is) **this**?’, and the San Antonio Aguas Calientes Kaqchikel interrogative ‘who?’ was given as *anchku’x la*’, literally ‘who (is/are) **that**?’. Finally, recall the San Marcos la Laguna Kaqchikel forms *naxwäch/naqwäch* ‘what?’ and *najwäch* ‘who?’, which also go back to a clausal construction roughly meaning something like ‘what/who is he/she/it?’. The final vowel *a* in forms such as Sololá Kaqchikel *achika* ‘who?, what?’ may also be attributed to the fused distal demonstrative *la*’.

should be analyzed as *t-e q-u'n* 3SG.GEN-RN 1PL.GEN-RN.²⁴ Judging on the data from some other Mam varieties, *t-e q-u'n* may just as well have the instrumental non-human meaning ‘with what?’. Thus, next to the San Pedro Necta form *te qu'n* Pérez *et al.* (2000:109) give forms such as Cajolá Mam *al q-u'n*, which are similarly translated only as ‘by whom?’, but on page 264 the same interrogative construction is used with the instrumental meaning ‘with what?’ (25).

Cajolá Mam (Mayan, Mamean; Guatemala; Pérez *et al.* 2000:264)

- (25) *al q-u'n x-Ø-ku'b'*
 IPW 3SG.GEN-RN RECENT.COMPL.DEP-3SG.OBJ-DIRECTIONAL
t-pa'-n-a' k'waal kw'il?
 3SG.A-break-AFF-MOVEMENT child pot
 ‘With what did the child break the pot?’

The interrogative ‘what?’ in the Mam varieties other than San Pedro Necta, usually has the form *ti()* and the like, e.g. *ti(i)*, *titi*, *tiri*, *tijiil*. In Cabricán Mam, ‘what?’ is *qal*, in all probability resulting from a truncation of something like San Ildefonso Ixtahuacán Mam *ti()* *t-qal* (lit.: what 3SG.GEN-RN) ‘with what?’ or *t-u'n t-qal* (lit.: 3SG.GEN-RN 3SG.GEN-RN) ‘with what?, by whom?’. In some varieties, the human and non-human interrogatives may become frozen with the particle/conjunction *tzun* ‘and, then, if’, sometimes only in certain syntactic functions, as in Santiago Chimaltenango Mam the interrogative *alkye* ‘who? (S, A, OBJ)’ in S-function is often *alkyetzun* (cf., e.g., Pérez *et al.* 2000:256-257).

The forms cognate to the San Pedro Necta Mam interrogative *alche(e)* ‘who?, what?’ in other Mam varieties typically mean ‘who?’: *a(a')l*, *alkye(e)*, *alke*, *elke*, *alqe(e)*, *elqe(e)*, *ankyee*, *anke*, *nkye*, *a()'loke*, *aloo*, etc. (Pérez *et al.* 2000:109-110, 254-268). This interrogative strongly resembles the Kaqchikel ‘who?, what?’ interrogative *a(n)chike* (and the like; cf. Section III.7.5.2) and in all probability has a similar origin as the latter, viz. a construction meaning ‘which one? (person or thing)’. However, the major difference between Mam and Kaqchikel is that in most Mam varieties this selective interrogative has developed into an exclusively human interrogative ‘who?’ and only in one

²⁴ Such constructions with two relational nouns for the “instrumental” interrogative pronominal meaning ‘with what?’ and sometimes ‘by whom?’ are common in some varieties of Mam. Compare, for instance, the San Ildefonso Ixtahuacán Mam constructions meaning ‘with what?, by whom?’: *t-u'n t-qal* 3SG.GEN-RN 3SG.GEN-RN and *al t-u'n* IPW 3SG.GEN-RN (Pérez *et al.* 2000:109, 264). As to the combination *t-e* 3SG.GEN-RN, in most varieties of Mam it is used to introduce the arguments encoding recipients, benefactors, sometimes patients, as well as focalized subjects of transitive verbs (Pérez *et al.* 2000:197, 202-205).

variety it has developed into a ‘who?, what?’ interrogative form, while in Kaqchikel the situation is practically the reverse.

As far as the form of the interrogatives *alche(e)*, *alqe(e)* and the like in pre-Mam is concerned the following can be said. To begin with, given that $k > ky > ch$ is a regular phonological development among the Mam dialects, while q normally remains unchanged (cf. Pérez *et al.* 2000:29-31, 35), this interrogative should be reconstructed with two variants of the final syllable *-ke(e)* and *-qe(e)*, where *-e(e)* is a relational noun and k - must be the 3PL.GEN prefix and q - the 1PL.GEN prefix. Recall, in this respect the discussion about *chi-k-e* ‘at/to them’ and *chi-q-e/cha-q-e* ‘at/to us’ as parts of the interrogative pronominal constructions with a selective meaning in Kaqchikel (Section III.7.5.2) and Tz’utujil (Section III.7.5.1). As to the variation between l and n , in some forms the lateral seems to be primary. Thus, first of all, the forms with l are much more frequent. Second, in some varieties with n , where the following velar stop can occasionally be dropped, we get l in its place. Thus, Comitancillo Mam interrogative *(a)nkye* ‘who? (S, A, OBJ)’ may become *al-tzun* in OBJ-function (cf. Pérez *et al.* 2000:109, 255-256), where *tzun* is a particle/conjunction meaning ‘and, then, if’. This l is most likely related to the Mam proximal demonstrative roots *la/lo/lu* and the similarly looking “dubitative particle” *la/lo*. The latter is used to introduce polar questions in combination with the preceding *(aa-)pa-/ (aa-)pi-* DEM-PQ-, *aa-tzu-* DEM-then-, *t-ee-pa/ t-ee-pi* 3SG.GEN-RN-PQ-, *t-ee-tzu-* 3SG.GEN-RN-then- (and the like) (cf. Pérez *et al.* 2000:99-100, 246-253). Kaufman (2003:1510, 1514-1515) reconstructs the Proto Mayan proximal root **lu*’ (although in Proto Yucatecan the same form is reconstructed with a distal meaning) and the Proto Quichean (proper, including Poqom languages) distal root **la*’. This hypothesis explains interrogative forms such as *a(‘)loke* or *aloo*’. The initial *a(a)*’ in the Mam interrogatives *alche* and the like must be the presentative/identificational deictic root ‘here/there/it is [X]’, which is the same *aa-* as in *(aa-)pa-/ (aa-)pi-* DEM-PQ-. In Mam and other Mayan languages, the same deictic element is used to form the independent personal pronouns and demonstratives and to introduce certain clefted arguments (e.g., for Mam see Pérez *et al.* 2000:98-101, 204-210). For Proto Mayan, this element is reconstructed by Kaufman (2003:1534-1535) as **ha*’.²⁵ In modern languages, the initial **h-* is often deleted or realized as $j/x/$, the glottal stop may be dropped and the vowel lengthened. The form **ha*’ would also explain the glottal stop in such interrogative forms as *a(a’)l* and *a(‘)loke* in some Mam varieties.

Summing up, for pre-Mam, the selective interrogative pronominal construction ‘which one? (person or thing)’ may be reconstructed as something like **[ha*’ ‘here/there/it is’ + DEM *la*’ (or *lu*) + *chi k-e(e)* ‘at/to 3PL.GEN-RN’ (or

²⁵ Kaufman (2003:1534-1535) also provides many reflexes; cf. the entry ‘he/she’ in Dienhart (1997) as well.

chi q-e(e) ‘at/to 1PL.GEN-RN’)]. Interestingly, this construction does not contain any element with an interrogative function. The interrogative pronominal meaning is expressed by the construction as a whole.

7.5.4 Uspanteka

Uspanteka/Uspantek/Uspanteco/Uspantec is a small Quichean Mayan language spoken by some 3000 people in the Uspantán municipality, in the Department El Quiché, Guatemala. Kaufman (1976, via Dienhart 1997) reports the interrogative *né:n* ‘who?, what?’ for the “Uspantán Centro” dialect of Uspanteka. In the Uspanteka texts collected by Huff & Huff (1971), I have found this interrogative, spelled as *nen*, only as ‘what?’ (26a) and as ‘what/which [N]?’ (26b). There are no examples with human interrogatives in Huff & Huff (1971).

San Miguel Uspantán Uspanteka (Mayan, Quichean; Guatemala; Huff & Huff 1971:504)

- (26) a. *i nen tatok niri?*
 and IPW you.are.looking.for here
 ‘And what are you looking for here?’
- b. *saber nen módo xkan li raloma?*
 who.knows IPW manner he.was on hillside
 ‘Who knows how (lit.: ‘what manner?’) he was on the hillside?’

The texts in Huff & Huff (1971) represent the variety spoken in the central town of the Uspanteka area, San Miguel Uspantán (cf. Huff & Huff 1971:241), which apparently is the same variety as Kaufman’s “Uspantán Centro”. It should be pointed out, however, that another source, Stoll (1884 via Dienhart 1997), gives for San Miguel Uspantán Uspanteka dedicated human and non-human interrogatives, viz. *ni* ‘who?’ and *le* ‘what?’.

The Uspanteka interrogatives *ni* ‘who?’ and *nen* ‘who?, what?’ strongly resemble Q’eqchi’ *ani* ‘who?’ (Quichean; Guatemala; Dienhart 1997) and the first part of Poqomchi’ *han wach* ‘who?’ (Quichean; Guatemala; Dienhart 1997). The word *wach* in the Poqomchi’ interrogative is clearly the same *wach* ‘face, surface, character, appearance, being, type, kind’, as in (18b) in Tz’utujil (Section III.7.5.1) and in the San Marcos la Laguna Kaqchikel forms *naxwäch/naqwäch* ‘what?’ and *najwäch* ‘who?’ (Section III.7.5.2). In turn, the forms *ani*, *hani* and *ni/nen* resemble the initial part of the pre-Kaqchikel-Tz’utujil interrogative **a(n)chinaq* ‘who?, what?, which/what [N]?’ (Section III.7.5.2). Therefore, it seems reasonable to suggest that the interrogative forms *ani*, *hani*, *ni/nen* and **a(n)chinaq* are due to the truncation of an earlier form

**hanichinaq* (or the like) ‘who?, what?, which/what [N]?’.²⁶ The final *n* of the Uspanteka form *nen* may be due to the reinforcement with some deictic root, probably a reflex of the Proto Mayan deictic root **ʔin*, which sometimes gives *ni/ne* (and the like) in the modern languages (cf. Kaufman 2003:1515, 1534), as in Uspanteka *niri* ‘here’ in (26a). In this respect, recall also the similar use of a demonstrative suggested in Section III.7.5.2 (footnote 23). to account for the final glottal stop in some Kaqchikel interrogative pronominals, such as *chikeʔ*.

7.5.5 Tzotzil

Tzotzil is a big Tzeltalan Mayan language, spoken in central Chiapas, Mexico. The Ethnologue distinguishes six Tzotzil languages: Chamula Tzotzil, Chenalhó Tzotzil, Huixtán Tzotzil, San Andrés Larrainzar Tzotzil, Venustiano Carranza Tzotzil, and Zinacantán Tzotzil. At least one variety of Tzotzil appears to have a single ‘who?, what?’ interrogative. Thus, Dienhart (1997) cites *buchʔu* ‘who?, what?’ from Cowan (1956) and *buchú* ‘who?, what?’ from Stoll (1884). Dienhart (1997) has not been able to identify the Tzotzil variety described in Stoll (1884). For Cowan (1956), Dienhart (1997) suggests the San Andrés Larrainzar municipality, but with a question mark. This is supported by the statement in Cowan & Merrifield (1968:284) that before 1958 Cowan did field work on Tzotzil only in the San Andrés Larrainzar municipality. It should be mentioned, though, that Delgaty (1964), another source cited by Dienhart (1997) for San Andrés Larrainzar, as well as Bochil and Zinacantán Tzotzil varieties, provides different human and non-human interrogative pronominals. However, given the great degree of dialectal variation usual for the Mayan languages, the discrepancy between the two sources does not look implausible.

In other Tzotzil varieties, the interrogative *buchʔu/muchʔu* means ‘who?’, while the non-human interrogative pronominal usually has the form *kʔu(sí)* ‘what?’. The interrogative *buchʔu* ‘who?, what?’ seems to be related to the interrogative *bu* ‘where?’ (Cowan 1956 via Dienhart 1997). The second syllable probably results from the contraction of an earlier construction. This construction may have been structurally similar to the Tzʔutujil phrase *chi-k-e* ‘at/to them’, *chi-q-e* ‘at/to us’, which is involved in the selective interrogative ‘which one?’ (cf. Section III.7.5.1), but with a different relational noun (instead of the relation noun *-e(e)*), containing *u* and a glottal stop, such as the Achiʔ *-uʔ* ‘with’ (Quichean; Guatemala; Shaw & Neuenswander 1966:32-33).²⁷ In other words, *buchʔu* may have originally been a selective interrogative ‘which one? (person or

²⁶ Note that this is a provisional reconstruction on the basis of the data only from these particular Quichean languages. It will be further developed in Section III.7.5.6.

²⁷ The final part of the human interrogative *an(ch)kuʔ(u)x* ‘who?’ in some varieties of Kaqchikel (Section III.7.5.2) might have the same origin as *-chʔu* of *buchʔu* in Tzotzil.

thing)’ that in most Tzotzil dialects evolved into the human interrogative ‘who?’ and only in one variety into the interrogative ‘who?, what?’.

7.5.6 *Mayan languages: concluding remarks*

Some five Mayan languages have been found to allow for the lack of differentiation between ‘who?’ and ‘what?’. Four of them, Tz’utujil, Kaqchikel, Uspanteka and Mam, belong to the Quichean-Mamean (or Eastern) branch of the family and one, Tzotzil, to the Cholan-Tzeltalan branch. For some languages, only certain dialects appear to possess a ‘who?, what?’ interrogative. Of the thirty one Mayan languages (as presented in Figure 2), I have been able to check at least one dialect of all but two languages, Chicomuceltec and Mocho/Motocintlec. At least in Kaqchikel, Mam and Tzotzil, the lack of differentiation between ‘who?’ and ‘what?’ appears to be primarily due to the fact that their ‘who?, what?’ interrogatives have developed from an originally selective interrogative pronominal construction ‘which one? (person or thing)’. As to the remaining two languages, Tz’utujil and Uspanteka, although their ‘who?, what?’ interrogatives appear to be cognate to those of Kaqchikel and Mam (the latter being structurally more complex), the lack of differentiation between ‘who?’ and ‘what?’ in these languages rather appears to be due to some common patterns of the formation of interrogative pronominals in Mayan languages in general and in Quichean-Mamean languages in particular. In the rest of the present section I will further elucidate this point. In particular, I will argue that (i) in Mayan languages the interrogative pronominals tend to be constructed as main clauses of cleft sentences and subsequently truncated and contracted; (ii) the Proto Quichean and Proto Mamean interrogative constructions, which resulted in the Tz’utujil, Kaqchikel, Uspanteka and Mam ‘who?, what?’ interrogatives respectively, were built without an interrogative pronominal, and were as such indifferent to the human vs. non-human distinction; (iii) the two interrogative constructions reconstructible for Proto Quichean and Proto Mamean share the same structure and differ only in its filling.

To begin with, it may be worth pointing out that no interrogative pronominal appears to have been convincingly reconstructed for Proto Mayan yet.²⁸ This is

²⁸ Thus, Kaufman (2003:1516-1517) gives only lower-level reconstructions of some interrogatives and in a few cases he just lists the forms of the interrogatives sharing the same meaning in various modern languages without suggesting a reconstruction. Patal Majzul *et al.* (2000:189) cites a Proto Mayan reconstruction **tu* ‘what?’ and nothing (“*---”) for ‘who?’. However, even this reconstruction of ‘what?’ looks rather dubious, given that for instance according to the data in Dienhart’s (1997) Mayan comparative dictionary, there are hardly any Mayan languages where the interrogative ‘what?’ at least contains a syllable *tu*. In some

remarkable, given that Mayan languages are among the best studied indigenous languages of the Americas, also with a long and successful tradition of diachronic research. The difficulty in reconstructing the Mayan interrogative pronominals is largely due to the fact that interrogative pronominals vary greatly not only between the different Mayan languages, but also between the dialects of a single language and sometimes even within one dialect. Remarkably, many other parts of the lexicon appear to be much more stable. This astonishing degree of variation in the interrogatives may have already become apparent in the discussion in the preceding sections. However, it may be instructive to demonstrate it here once more with (27) for the differences between languages and with (28-29) for the differences between and within dialects.

The interrogative 'who?' in the Guatemalan Mayan languages (from the comparative wordlist in Mayers 1966:275)

(27)	Achi'	<i>pachinoq</i>
	K'iche'	<i>jachin</i>
	Kaqchikel	<i>achike</i>
	Poqomchi'	<i>ha wach</i>
	Poqomam	<i>hayo'</i>
	Q'eqchi'	<i>ani</i>
	Ixil	<i>ab'il</i>
	Awakateka	<i>na'j</i>
	Mam	<i>'alkyee</i>
	Jakalteka	<i>mak</i>
	Chuj	<i>mach</i>
	Ch'orti'	<i>chi</i>

The interrogative 'with what? (instrument), how?' in some K'iche' dialects (Par Sapón & Can Pixabaj 2000:94, 206)

(28)	Santa Lucía Utatlán	<i>jasu'wach</i>
	Cantel	<i>jas uwuch</i>
	Santa María Chiquimula	<i>jas (su'), sur</i>
	Santa Cruz de El Quiché	<i>su'</i>
	Rabinal	<i>sa'</i>
	San Antonio Ilotenango	<i>jas(o)</i>
	Totonicapán	<i>jas</i>
	Cubulco	<i>wach</i>

dialects of Kaqchikel, 'what?' may be expressed by something like *atux*, (cf. Table 4 in Section III.7.5.2) and in Ch'orti' 'what?' is *tuk'a*.

The interrogative ‘who?’ in some K’iche’ dialects (Par Sapón & Can Pixabaj 2000:95, 203)

(29)	Totonicapán	<i>jachin(aq)</i>
	Samayac	<i>(ja)chin</i>
	Zunil	<i>qachi(n)</i>
	Cubulco	<i>chinoq, (chi)na, pa</i>
	Cunén	<i>naq</i>

As can be observed from the examples in (28-29), interrogatives in Mayan languages appear to be particularly prone to augmentation and subsequent truncation and contraction. Thus, the element *wach* in (28) is clearly the same *wach* ‘face, surface, character, appearance, being, type, kind’, as in (18b) in Tz’utujil (Section III.7.5.1), reproduced here as (30). Recall also the San Marcos la Laguna Kaqchikel forms *naxwäch/naqwäch* ‘what?’ and *najwäch* ‘who?’ (Section III.7.5.2) and Poqomchi’ *han wach* ‘who?’ (Section III.7.5.3).

San Juan la Laguna Tz’utujil (Mayan, Quichean; Guatemala)

- (30) *naq r-wach?*
 IPW 3SG.GEN-face
 ‘Who/what is it?’ (Dayley 1985:237)

It appears that almost every part of the interrogative can be truncated (or contracted). In this respect, recall as well the Kaqchikel interrogative construction *achike (r-)uma* ‘why? (lit.: IPW (3SG.GEN-)cause)’ that can be contracted to *aruma* ‘why?’, where in turn the initial element *achike* itself must go back to another complex construction **a(n)chinaq chike* (or the like; cf. Section III.7.5.2). As a result of all these augmentation and truncation processes, cognate forms may become unrecognizable as such, as for instance Zunil K’iche’ *qachi* and Cunén K’iche’ *naq* ‘who?’ in (29) or Rabinal K’iche’ *sa’* and Cubulco K’iche’ *wach* ‘with what? (instrument), how?’. Furthermore, we may get two (almost) identical forms with different meanings in two different idioms, e.g. Cubulco K’iche’ *wach* ‘with what? (instrument), how?’ vs. Purulhá Poqomchi’ *awach* ‘who?’ (Malchic Nicolás *et al.* 2000:92) or Cunén K’iche’ *naq* ‘who?’ vs. San Juan la Laguna Tz’utujil *naq* ‘who?, what?’.

Why and where exactly augmentation and especially truncation take place in a particular idiom is clearly a complex issue conditioned by many factors. To name just a few, one may think about (i) the tendency to shorten frequently used items, (ii) the variation in stress patterns, (iii) the relevance of the particular element for the semantics of the whole construction, (iv) the tendency for reinforcement due to the emphatic, focalized nature of the interrogatives, (v) the tendency for semantic specification, etc.

Another conclusion that is suggested by such an extensive formal variation in the forms of the interrogatives is that interrogative pronominals in Mayan languages must have a constructional origin. In fact, these constructions seem to be clauses, viz. main clauses of clefted sentences. Among other things, this may be supported by the following facts: (i) the interrogatives are sentence-initial and Mayan languages are generally predicate-initial, (ii) often, the interrogatives contain initial elements *(j/h)a-* (cf. Section III.7.5.3) and the same forms are also used to introduce clefts in many languages, (iii) in some languages, a special dependent form of the verb may need to be used after the interrogative pronominal just as after other clefted elements, as in the Cajolá Mam examples (31a, b), where the dependent “recent completive” aspect marker *x-/s-* is used instead of the corresponding independent marker *ma* (31c).

Cajolá Mam (Mayan, Mamean; Guatemala)

- (31) a. *alke s-Ø-aj chlee-’n*
 who RECENT.COMPL.DEP-3SG.OBJ-DIRECTIONAL hug-AFF
t-e qya?
 3SG.GEN-RN woman
 ‘Who has hugged the woman?’ (Pérez *et al.* 2000:254)
- b. *a Nikte’ x-Ø-xi’*
 it.is PROP RECENT.COMPL.DEP-3SG.OBJ-DIRECTIONAL
ooni-’n t-e Saqb’ech
 help-AFF 3SG.GEN-RN PROP
 ‘It is Nikte’ who has helped Saqb’ech’ (Pérez *et al.* 2000:168)
- c. *ma Ø-txi’ t-ooni-’n*
 RECENT.COMPL.INDEP 3SG.OBJ-DIRECTIONAL 3SG.A-help-AFF
Nikte’ Saqb’ech
 PROP PROP
 ‘Nikte’ has helped Saqb’ech’ (Pérez *et al.* 2000:167-168)

To a large extent, the degree of flexibility of the resulting interrogatives must depend on the kinds of elements that are used to build the original interrogative constructions. Thus, various deictics, relational nouns, conjunctions and discourse particles, i.e. elements forming sets of semantically comparable and largely exchangeable items, appear to be particularly suitable for this role. On the contrary, nouns with meanings such as ‘person’, ‘man’ or ‘thing’ would be much less suitable. In this respect, recall for instance the interrogative construction ‘which one?’ *[*ha* ‘here/there/it is’ + DEM *la*’ (or *lu*) + *chi k-e(e)* ‘at/to 3PL.GEN-RN’ (or *chi q-e(e)* ‘at/to 1PL.GEN-RN)] reconstructed in Section III.7.5.3 for pre-Mam. It contains only deictics, a personal affix, a relational noun and an adposition.

If we now compare this construction to the interrogative **hanichinaq* (or the like) ‘who?, what?, which/what [N]?’ provisionally reconstructed in Section III.7.5.4 on the basis of the data of the Quichean languages Uspanteka, Q’eqchi’, Poqomchi’, Tz’utujil and Kaqchikel, we may note the similarity in their initial parts, viz. **ha’* and **ha-*. What is more, just as *la’* (or *lu’*) is a deictic in the pre-Mam construction, the *ni* of **hanichinaq* may be a deictic root as well, viz. a reflex of the Proto Mayan deictic root **in*, which sometimes gives *ni/ne* (and the like) in the modern languages (cf. Kaufman 2003:1515, 1534; cf. also Section III.7.5.4). Thus, instead of **hanichinaq* we should rather reconstruct something like *[*ha’* ‘here/there/it is’ + *in* DEM + *chinaq* (or the like)].²⁹

The forms of the interrogative pronominals in several other Quichean languages indicate the possibility of further variations on the basis of the same construction *[*ha’* ‘here/there/it is’ + DEM + *chinaq* (or the like)]. Thus, forms such as Achi’ *pachin(oq)* ‘who?’ and *pachin ike* ‘which of them?’ (Shaw & Neuenswander 1966:33) suggest that the interrogative construction at issue may also be *[*ha’* + *pa* + (DEM) + *chinaq* (or the like)], thus resembling the construction *aa-pa-la* used in Mam to introduce polar questions (cf. Section III.7.5.3). The element *pa* is also identical to the adposition *pa* ‘in, to’ in Achi’ and many other Mayan languages. It also shows up recurrently (as the first element) in the interrogatives meaning ‘where?’ in Mayan languages.³⁰ Another possible variation is suggested by several Quichean interrogatives with initial *qa-* instead of (*h/j*)*a-*, such as San Juan Sacatepéquez Kaqchikel *qaxchiké* ‘who?’ (Stoll 1884 via Dienhart 1997), Palín Poqomam *qa’keh* ‘who?, which one (person or thing)?’, *qa’sa’* ‘what?, how?’ (Malchic Nicolás *et al.* 2000:92) and Zunil K’iche’ *qachi(n)* ‘who?’. In particular, it seems that we are dealing here with the same element *qa(s)* that begins the phrase meaning ‘it is so, it is true, yes’ in many Quichean-Mamean Mayan languages, as in *qa tziij* ‘it is true’ (*tziij*

²⁹ The possible meaning of the element *chinaq* (or the like) will be discussed later in this section.

³⁰ For instance, in Achi’ the locative interrogative is *pa ... wi* ‘to/in where?’ or *chi ... wi* ‘to/at where?’ (Shaw & Neuenswander 1966:39), where *wi* looks like a demonstrative root in origin (compare the Tz’utujil postverbal emphatic particle *wi’*, which is used when the oblique argument, such as locative adverbs, prepositional or relational noun phrases of certain kinds, is fronted “because [it is] questioned, contrastive, or emphatic”, Dayley 1985:256). In Palín Poqomam, ‘where?’ is *pa re’*, where *re’* is a demonstrative as well, whereas in Ribalcó Poqomchi’ ‘where?’ is *pila’* (cf. Malchic Nicolás *et al.* 2000:92-94), comparable to the Mam polar question introducer (*aa-*)*pi-la*/(*aa-*)*pa-la* mentioned above. In some western varieties of K’iche’, the interrogative ‘where?’ can be found as *pa wi* (Rabinal K’iche’), *pa chi* (San Miguel Chicaj K’iche’), *pa* (Cubulco K’iche’) (Par Sapón & Can Pixabaj 2000:94). In several varieties of Kaqchikel, ‘where?’ can be found as *ape(’)* and *pache’* (Patal Majzul *et al.* 2000:74).

is a noun meaning ‘word’) in Rabinal, San Miguel Chijac and San José Chiquilajá K’iche’ varieties (Par Sapón & Can Pixabaj 2000:95; cf. also Kaufman 2003:733-734 for more examples). In San Juan la Laguna Tz’utujil, the “adverbial particle *qas* ‘very, really, a lot’ [...] very often occur[s with] predicate adjectives”, as in *qas at nim* ‘you are really big’ (lit.: be.really 2SG.S big), which “seems to be to unambiguously mark adjectives as predicates, as opposed to modifiers”; what is more, *qas* “only occurs with predicates in Tzutujil, never with modifiers” (Dayley 1985:201). There is no need to presume, however, that in interrogatives such as *qaxchiké*, *qa’keh*, and the like, this *qa(s)* was used instead of **ha’* ‘here/there/it is [X]’. It could have been used in combination with **ha’* just as well (probably, similarly to *pa* discussed above).

If we leave aside the variations just discussed, it appears to be possible to reconstruct for Proto Quichean the general interrogative pronominal construction **[ha’* ‘here/there/it is’ + DEM (e.g., *’in*) + something like *chinaq*] ‘who?, what?’, which was freely expandable into the selective interrogative construction ‘which one?’ through the use of phrases structurally similar to Tz’utujil *chi k-e* ‘at/to 3SG.GEN-RN’. The element *chinaq* (or the like) in the interrogative construction **[ha’* + DEM + *chinaq* (or the like)] is particularly interesting because it seems to be indifferent to the human vs. non-human distinction and because it does not seem to be related to interrogative pronominals in the other branches of the Mayan family. In fact, it seems that in origin it is not an interrogative at all. Instead, *chinaq* (and the like) looks very much like the resultative (or perfective) participle or third person singular perfect form of a reflex of the Proto “Central Mayan”³¹ intransitive verb **kih* ‘to say’ (Kaufman 2003:739), which in modern languages is typically used as a “quotative” verb and is often irregular.³² In other words, the interrogative construction at issue probably had the form **[ha’* ‘here/there/it is’ + DEM + *Ø-kih-naq* 3SG.S-say-PRF/PTCP] and literally meant something like ‘this/that that [P] is which is said “...”’, implying ‘this/that that [P] is who/what?’. Such an interrogative pronominal construction without, as it were, an interrogative pronominal would be comparable to the Austronesian ‘who?, what?’ interrogatives derived from the word ‘name’ discussed in Section

³¹ That is, the common ancestor of the Eastern Mayan (i.e., Quichean-Mamean) and Western Mayan (i.e., Cholan-Tzeltalan and Kanjobalan-Chujean) branches.

³² Thus, Tz’utujil has a “highly irregular intransitive verb [...] *che’*- (~ *chi*- ~ *e’*- ~ *i*-) ‘say’, which is used in quoting someone directly [and] is etymologically related to the quotative particle *cha*”, that “may be translated as ‘he/she said’, ‘they say’, ‘it is said’” (Dayley 1985:108, 260). The resultative/perfective participle (or perfect) form of this verb in Tz’utujil is *che’-naq*. The suffix *-naq* (and the like) is the regular resultative (or perfective) participle/perfect marker with intransitive verbs in Quichean-Mamean languages, e.g. Tz’utujil *-naq* (Dayley 1985:77-79), Kaqchikel *-(i)näq* (Patal Majzul *et al.* 2000:81) or Mam *-na* or (depending on the dialect) *-ni* (Pérez *et al.* 2000:119).

III.4.2.2.1.1 and especially the ‘who?, what?’ interrogatives in some non-North Arawakan languages (cf. Section III.8.2).

The proposed Proto Quichean reconstruction *[*ha* ‘here/there/it is’ + DEM + Ø-*kih-naq* 3SG.S-say-PRF/PTCP] appears to be indirectly supported by the data from the Mamean branch of the Mayan family. The support comes from the Mamean languages other than Mam, since the pre-Mam reconstruction *[*ha* ‘here/there/it is’ + DEM *la*’ (or *lu*) + *chi k-e(e)* ‘at/to 3PL.GEN-RN’ (or *chi q-e(e)* ‘at/to 1PL.GEN-RN)] ‘which one?’ advanced in Section III.7.5.3 appears to lack the part comparable to Ø-*kih-naq* in the Proto Quichean reconstruction. However, more information can be deduced from the following interrogative pronominals in the remaining Mamean languages: Ixil *abil/ja’l* ‘who?’, Tektiteka/Teco *aabl/’aabil* ‘who?’, Awakateka/Aguateco *mbii’/jal* ‘what?’ (Dienhart 1997). The element *bii’/bi/b* in these interrogatives seems to be related to the Proto Mayan word **b’ih* ‘name’ and its derivative, the transitive verb **b’ih.i* ‘to name, to say, to tell something’ (cf. Kaufman 2003:737-738; cf. also Dayley 1985:110 on the respective San Juan la Laguna Tz’utujil reflexes). The final *-l* must be the same element as the “impersonal passive” suffix *-l/-j* in Mam,³³ as in (32).

Mam (Mayan, Mamean; Guatemala; Pérez *et al.* 2000:223-224)³⁴

- (32) a. *n-chi* *tzuy-I*
 INCOMPL-2PL.S take-PASS
 ‘You are being hold (?detained)’, ‘They are holding you (PL)’ (the original Spanish translation: ‘Están siendo agarrados’ and ‘Los están agarrando’)
- b. *n-Ø-joy-I* *pwaq*
 INCOMPL-3SG.S-look.for-PASS money
 ‘They are looking for the money’ (the original Spanish translation: ‘Se busca dinero’ and ‘Están buscando dinero’)
- c. *n-Ø-q’uum-j*
 INCOMPL-3SG.S-say-PASS
 ‘It is being said’, ‘They are saying it’ (the original Spanish translation: ‘Se está diciendo’ and ‘Lo están diciendo’)

Thus, in all probability, in Proto Mamean the general interrogative construction ‘who?, what?, which one?’ had the form *[*ha* ‘here/there/it is’ + DEM + *n-Ø-*

³³ In Mam, the “passive” marker *-l/-j* is used only with transitive verbs having third person agents (“se aceptan agentes oblicuos únicamente en terceras personas”, Pérez *et al.* 2000:122, 223). Furthermore, “se puede comprobar que el significando que resulta puede ser equivalente a la de una forma impersonalizada, un tanto menos que pasiva” (Pérez *et al.* 2000:223).

³⁴ The dialect was not specified in the source, because the examples are supposed to illustrate the feature common to all Mam dialects.

b'ihí-1 INCOMPL-3SG.S-name-PASS] and literally meant something like ‘this/that that [P] is which is called “...”?’’, implying ‘this/that that [P] is who/what?’.

The fact that the same structure but with somewhat different filling can be reconstructed for the interrogative pronominals of two closely related branches of the Mayan family, viz. Quichean and Mamean, further strengthens the plausibility of the respective reconstructions.

7.6 Garifuna and Island Carib

Garifuna/Garífuna, also known as Central American or Black Carib, is a Northern Arawakan language spoken primarily on the Atlantic coast of Belize, Guatemala and Honduras. Garifuna speakers are descendants of some five thousand Island Caribs who were deported in 1797 by the British from the island of St. Vincent in the Lesser Antilles to the Roatán Island just to the north off the coast of present-day Honduras (Taylor 1958b:153). From there they have later spread to the territory they currently inhabit. The closest Arawakan relative of the Garifuna is the Island Carib (or Iñeri) language that was still spoken until around 1920 on the islands of St. Vincent and Dominica. However, only the Dominican variety had been recorded (Taylor 1958b:153-154; cf. also Taylor 1957 for a sketch and references to the original sources).

Already in the pre-Columbian times, the Arawakan ancestors of the Island Caribs (and ultimately Garifuna) were conquered by some Carib speaking group. An interesting kind of diglossia had developed in the Island Carib community: the men used some kind of Carib-Arawakan pidgin and the original Carib language of the conquerors, while women and children largely continued to use the original Arawakan idiom (cf. Aikhenvald 1999:74-75; Fleming 1966:303-305; Taylor & Hoff 1980). Eventually, the special men’s speech fell into disuse, but the Island Carib underwent an important lexical and grammatical Carib influence. During the 17th and the 18th centuries the Island Carib community had incorporated an important group of “free and escaped African slaves who had sought refuge among the Island Carib Indians” (Fleming 1966:303), which is why the Garifuna have often been referred to as the *Black Caribs*.

According to the Ethnologue (cf. also Aikhenvald 1999:69), Island Carib and its offshoot Garifuna belong to the so-called Caribbean subgroup of the Arawakan languages, which also includes Lokono/Arawak (Suriname, French Guiana, Guyana), Wayuu/Guajiro (Colombia/Venezuela), Parajuanó/Añun (extinct or almost extinct; Venezuela), Taino (extinct; the Bahamas and the Greater Antilles), Caquetio (extinct; Venezuela), Shebayo (extinct; Trinidad Island). The bulk of the Arawakan languages is located in the inland South America, stretching as far south as the north of Argentina, and will be considered together with other South American languages in Section III.8.2.

Both Dominican Island Carib and Garifuna appear to use general ‘who?, what?’ interrogatives. The forms of the interrogatives are also very similar. Thus, Dominican Island Carib has *kát(a)-* (Taylor 1956b:140, 1957:299-300).³⁵ The Hopkins (Belize) dialect of Garifuna, as described by Taylor (1951a-b, 1955, 1956a-b, 1958a-b) has two forms, one invariable, *ka(-)* ‘who?, what?’, and one inflected for gender-number-person, *káta-* ‘be who?, be what?’. The Livingston (Guatemala) dialect of Garifuna, as described in Fleming (1966), is reported to have *ka-* ‘who?, what?’ and *káta-* only ‘what?’. However, the last description provides only a wordlist and a few phrasal examples, which may suggest that the lack of the human meaning ‘who?’ for *káta-* in this source is accidental.

By way of illustration, let us examine the ‘who?, what?’ interrogatives of Hopkins Garifuna in more detail. The use of the invariable form *ka* ‘who?, what?, which one?, which/what [N]?’ is illustrated in (33-35).

Hopkins Garifuna (Northern Arawakan, Caribbean; Belize)³⁶

- (33) a. *ka sa l-adógo-ba-ɥ?*
 IPW PTCL 3SG.M-make.NMLZ-SUB-3SG.F
 ‘What did he make?’ (lit.: ‘What is in fact she that he made?’) (Taylor 1956a:6)
- b. *ka sa l-úma-ri-ba-ɥ?*
 IPW PTCL 3SG.M-be.with-NMLZ-SUB-3SG.F
 ‘Who is his wife?’ (Taylor 1956b:142)
- c. *ka-ba l-áfara?*
 IPW-IPFV 3SG.M-beat
 ‘Whom will he beat?’ (Taylor 1951a:30)
- d. *ka-ba áfara-i?*
 IPW-IPFV beat-3SG.M
 ‘Who will beat him?’ (Taylor 1951a:30)
- (34) *ká siuámaj bu-bá-ɥ há-dageɛ?*
 IPW agreeable 2SG-be.IPFV-3SG.F 3PL.AN-among.from
 ‘Which (female) one of them (animate) do you like?’ (Taylor 1956b:147)
- (35) a. *ká dóbu?*
 IPW stone
 ‘What stone?’ (Taylor 1956a:5)

³⁵ Taylor (1956b, 1957) provides only examples of *kát(a)-* used as ‘who?’, but the existence of its second non-human meaning ‘what?’ can be deduced from the discussion in the respective sources.

³⁶ The acute accent in the examples marks stress, which is not fixed in Garifuna.

- c. *ká i-dóbu-ri?*
 IPW POSS-stone-POSS
 ‘Whose stone?’ (Taylor 1956a:5)
- d. *ka(-u) i-sábu-te lèa?*
 IPW(-to) POSS-soap[M]-POSS M.SG.this
 ‘Whose soap is this?’ (Taylor 1951b:225)
- d. *ká-uori?*
 IPW-machete
 ‘Whose machete?’ (Taylor 1956a:5)

The variable form *káta*- ‘be who?, be what?’ must be inflected for gender-number-person (36).

Hopkins Garifuna

- (36) a. *káta-i?*
 IPW-3SG.M
 ‘Who is he?’ or ‘What is it(M)?’ (Taylor 1956b:140-141)
- b. *káta-u?*
 IPW-3SG.F
 ‘Who is she?’ or ‘What is it(F)?’ (Taylor 1956b:140)
- c. *káta-ia b-agóburigu-bá-ia?*
 IPW-3PL.ANIM 2SG-parents-be-3PL.AN
 ‘Who are your parents?’ (Taylor 1956b:140)

In the singular, Garifuna has two genders, masculine and feminine, each containing both animate and inanimate nouns. In the plural, the gender distinction is neutralized, but the plural agreement pattern appears to be restricted to the animates. The use of the masculine agreement pattern in sentences like (37), where the gender of the referent is unidentifiable, or (38), where the controller is an action nominalization, seems to suggest that the masculine is the functionally unmarked agreement pattern in Garifuna.

Hopkins Garifuna

- (37) *káta-i gía?*
 IPW-3SG.M PTCL
 ‘What is it?’, ‘What can it be?’, ‘What’s the matter?’, ‘What do you want?’ (Taylor 1956b:140, 142)
- (38) *I-ubáragie t-erégu...*
 3SG.M-before 3SG.F-speak.NMLZ
 ‘Before she spoke... (lit.: ‘before it, her speaking’)’ (Taylor 1958a:39)

In addition to their interrogative function, the singular forms *káta-i* (M) and *káta-ɥ* (F) can also be used as nouns meaning ‘thing’, as in *ábɔ́ káta-i ítara* ‘such a thing’ (*ábɔ́* ‘one’, *ítara* ‘such’), *káta-i léa* ‘this thing’ (*léa* ‘M.SG.this’), *káta-ɥ tóa* ‘this thing’ or ‘the male organ (a euphemism)’ (*tóa* ‘F.SG.this’), *sɥ káta-i* ‘everything’ (*sɥ* ‘all’), *ámu káta-i* ‘other things, another thing’ (*ámu* ‘(an)other’) (Taylor 1956b:140, 1958a:37). Furthermore, it seems that occasionally at least the non-interrogative *káta-* can be used without gender marking, as in *ní káta* ‘nothing’ (Taylor 1956b:146).

As will be demonstrated in Section III.8.2, the lack of differentiation between ‘who?’ and ‘what?’ is widespread among the Arawakan languages, but the forms of the interrogatives vary greatly from language to language. A detailed comparative study would be necessary to account for the details of this variation.

8 South America¹

8.1 Introduction

South America is a region with both one of the highest degrees of linguistic diversity in the world and one of the highest numbers of languages that are still in need of a detailed description. All in all, data for some 180 South American languages have been consulted. However, for the reasons explained in Section III.1, it happened to be impossible to examine the languages of South America in the same degree of detail as many of the Old World languages. Still, 20-25% of the South American languages in my sample have been found to allow a lack of differentiation between ‘who?’ and ‘what?’, which appears to be one the highest ratios in the world. Within South America, the highest concentration of languages allowing for the lack of differentiation between ‘who?’ and ‘what?’ can be found in southwestern Amazonia, more particularly in the region that can be roughly defined as the upper parts of the basins of Río Ucayali, Río Purús and, especially, of Río Madeira, cf. Map 1.

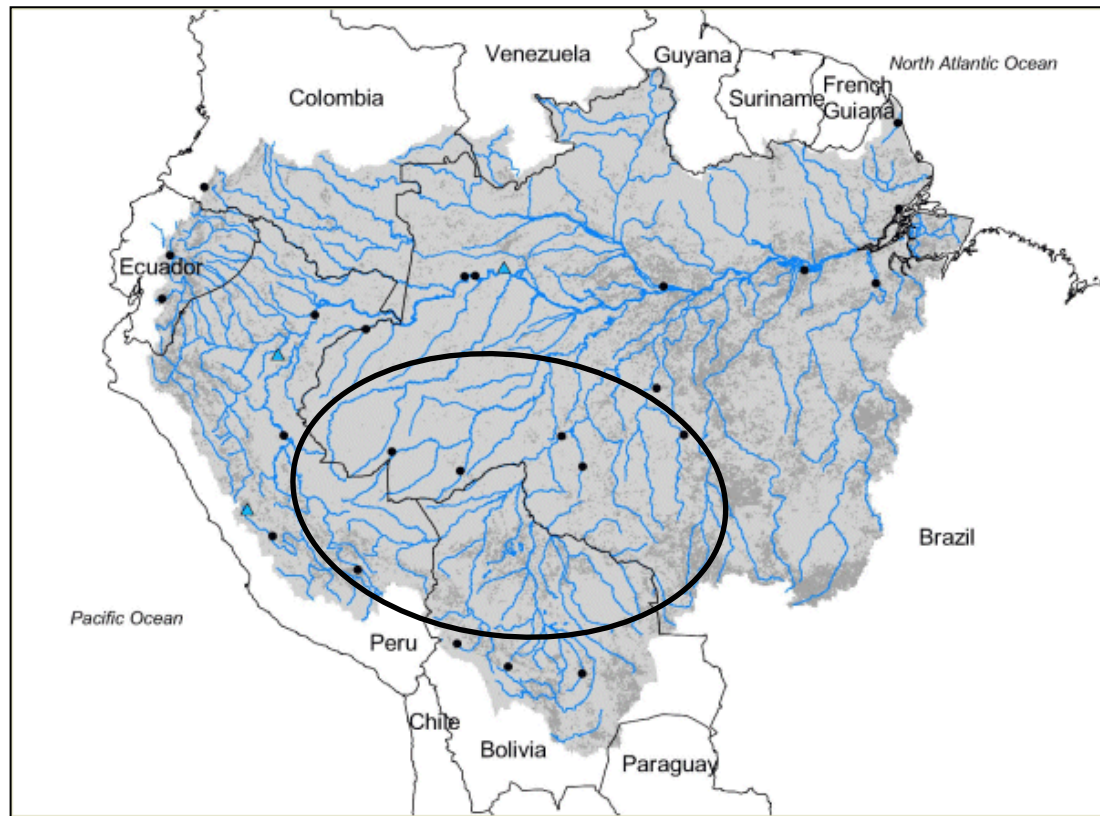
Of all the South American language families, Arawakan, which is the biggest and geographically the most extended family, is also the family with the largest number of languages with ‘who?, what?’ interrogatives. Therefore, I will begin by discussing Arawakan languages in Section III.8.2, proceeding then to the languages spoken in the northwestern parts of the continent and subsequently moving down to the south(east): Section III.8.3 Arhuaco/Ika (Chibchan; Colombia), Section III.8.4 Dâw (Makú; Brazil), Section III.8.5 Urarina (Isolate; Peru), Section III.8.6 Paumarí (Arauan; Brazil), Section III.8.7 Itene/More (Chapacura-Wanham; Bolivia), Section III.8.8 Tacanan languages, Section III.8.9 Tupí languages, Section III.8.10 Macro-Jê languages, Section III.8.11 Guaicuruan languages, and Section III.8.12 Matacoan languages.²

The attested areal distribution of the languages with ‘who?, what?’ interrogatives in South America largely coincides with the distribution of the Arawakan languages. In this respect, it may be interesting to mention that Arawakan-speaking societies of “wetland agriculturalists” were involved in extensive trade and intermarriage relations with other groups and were some of

¹ Including the islands of the Caribbean.

² There is one more language in my sample that may prove to allow a lack of differentiation between ‘who?’ and ‘what?’, viz. Sabanê(s), a Nambiquaran language spoken in the state of Mato Grosso, Brazil, and described by Antunes de Araujo (2004). Unfortunately, (due to typos?) the description of interrogatives in Antunes de Araujo (2004) is somewhat incoherent so that no clear conclusions can be drawn from it.

Map 1. A rough delimitation of the southwestern Amazonia, the region with the highest concentration of languages allowing a lack of differentiation between ‘who?’ and ‘what?’ in South America (based on the map of the Amazon’s watershed by the World Resources Institute, <http://www.wri.org>)



the most “powerful and expansive polities of pre-Columbian Amazonia” (Hornborg 2005). Furthermore, intensive linguistic contacts between the Arawakan-speaking societies and their neighbours took place as well. All this may be interpreted as suggesting that the presence of the Arawakan languages could have stimulated the development of ‘who?, what?’ interrogatives in other languages. However, one should not lose sight of the fact that there are much less indigenous languages (left) in the (north)eastern and southern parts of South America, which might be why the areal distribution under consideration is somewhat skewed in favour of the more remote and less easily accessible parts of the Amazonia. At the same time, there is an important element of this distribution that seems to remain unaffected by the reservations just pronounced. Thus, it may be noted that the lack of differentiation between ‘who?’ and ‘what?’ is virtually unattested in the languages of the Andes.

South America is not only the region with the highest number of languages allowing for the lack of differentiation between ‘who?’ and ‘what?’, it is also the region with some of the most typologically unusual systems of interrogatives, such as the ones found in Arawakan (Section III.8.2.1) and Chapacura-Wanham

languages (Section III.8.7). Perhaps the most interesting feature of these systems is that interrogatives appear to be built as clausal constructions without any interrogative or indefinite nominal elements, with verbal elements used instead. Also remarkable is that in South America interrogatives sometimes vary a lot even between closely related languages, even though the number of forms found in each particular language is at the same time very limited.

The lack of differentiation between ‘who?’ and ‘what?’ in the South American languages discussed may have various sources. At least the following possibilities seem to be attested: (i) the development of the locative ‘where?’ interrogatives into selective and subsequently non-selective interrogative pronominals (Arawakan, Guaicuruan, possibly also Arauan, Chapacura-Wanham, Macro-Jê), (ii) the origin of the respective interrogatives in clausal constructions, presumably based on verbs meaning ‘do, say, be’ and ‘name’ (Arawakan, possibly also Tacanan and Urarina), (iii) the peculiarities of the gender semantics and the form of gender markers (Arawakan, possibly also Arauan, Chapacura-Wanham, Guaicuruan and Matacoan), (iv) the conventionalization of a filler (placeholder) meaning ‘whatsi(t)sname’ into a full-fledged interrogative pronominal (possibly Macro-Jê and Tacanan).

8.2 Arawak(an) languages³

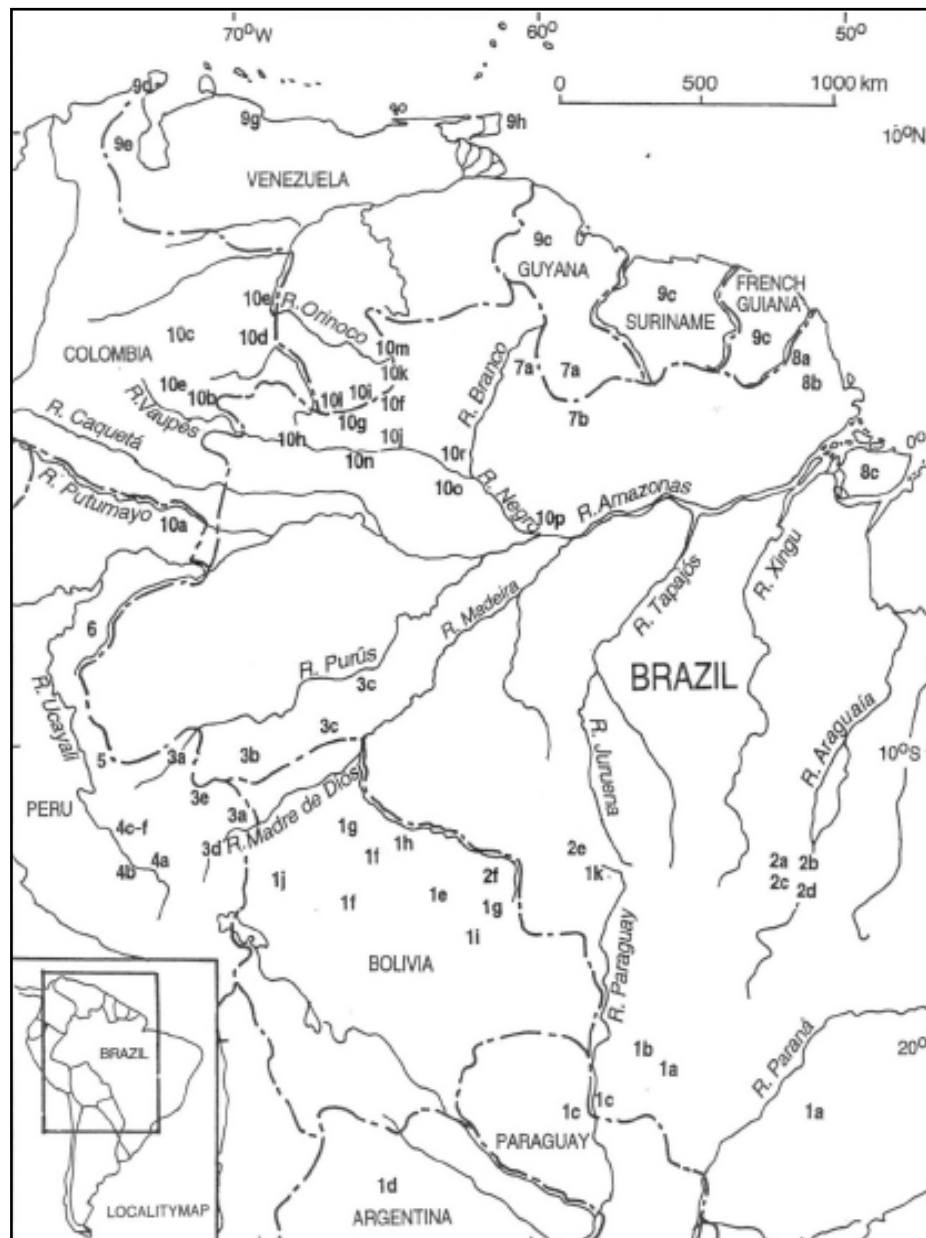
The Arawakan language family “contains the largest number of languages in South America”, viz. about 40 living languages and about 20 extinct languages “on which materials are available” (Aikhenvald 1999:65, 67-71).⁴ It is also geographically the most extended family in the region. Arawakan languages are or used to be spoken as far north as the Bahamas and Antilles and as far south as the north of Argentina. Since the end of the 18th century an Arawakan language Garifuna is also spoken in North America (currently in Belize, Honduras, Guatemala, and Nicaragua; cf. Section III.7.6). The existing internal classifications of the Arawakan languages are largely geographically based. According to Aikhenvald (1999:73-75), linguistically the main dividing line within the family lies between the North Arawakan languages, on the one hand, and non-North Arawakan languages, viz. South and Southwestern Arawakan, on

³ The family has been variously referred to in the literature as *Arawakan*, *Arawak*, *Aruák*, *Maipur(e)an*. However, these terms are not always identical as to the range of languages they are intended to cover (cf. Aikhenvald 1999:73 for a brief discussion). I use the term *Arawakan* in the same sense as Aikhenvald (1999) uses *Arawak*.

⁴ Much higher numbers, such as 89 or 154, may be found in some other sources (cf. Aikhenvald 1999:65 for the references). As is often the case elsewhere in the world, the distinction between a language and a dialect is a tricky issue, very much dependent on the parameters preferred.

the other. The Ethnologue suggests a more fractional subdivision between the Central, Eastern, Northern, Southern, and Western groups. The Map 2 provides the approximate locations of the Arawakan languages (without Garifuna and the long extinct languages of the Bahamas and Antilles) and presents “a cautious assessment of what [Aikhenvald (1999:65)] believe[s] to be distinct languages”.⁵

Map 2. Arawakan languages with approximate locations (without Garifuna and the extinct languages of the Caribbean; following Aikhenvald 1999:65-71)



⁵ The spellings of the language names used in this section may differ slightly from those used by Aikhenvald and reproduced here in the legend to the Map 2.

non-North Arawakan**South Arawakan****1 South Arawakan**

- 1a** Terêna
- 1b** †Kinikinao
- 1c** †Guané/Layana
- 1d** †Chané/Izoceño
- 1e** Bauré
- 1f** Moxo/Ignaciano
- 1g** Moxo/Trinitario
- 1h** †Paiconeca
- 1i** †Pauna
- 1j** †Apolista
- 1k** Salumã (Enawenê-nawê)

2 Pareci-Xingu

- | | |
|----------------------|-------------------------|
| <i>Xingu</i> | <i>Pareci-Saraveca</i> |
| 2a Waurá | 2e Pareci/Haliti |
| 2b Mehinaku | 2f †Saraveca |
| 2c Yawalapiti | |
| 2d †Kustenaú | |

Southwestern Arawakan**3 Piro-Apuriná**

- 3a** Piro (Maniteneri, Maxineri)
- 3b** Chontaquiro
- 3c** Apurina/Ipurina, Cangiti
- 3d** (†)Iñapari
- 3e** Mashko-Piro (?dialect of Iñapari)

4 Campa

- 4a** Ashaninca
- 4b** Asheninca
- 4c** Caquinte
- 4d** Machiguenga
- 4e** Nomatsiguenga
- 4f** Pajonal Campa

5 Amuesha**6 †Chamicuro****North Arawakan****7 Rio Branco**

- 7a** Wapishana
- 7b** Mayawana/Mapidian/?Mawakwa

8 Palikur

- 8a** Palikur
- 8b** †Marawan
- 8c** †Aruan/Aroã

9 Caribbean, or Extreme North

- 9a*** †Island Carib (Iñeri)
- 9b*** Garifuna (Black Carib)

TA-Arawakan subgroup of Caribbean

- 9c** Lokono/Arawak
- 9d** Guajiro/Wayyu
- 9e** Añun/Parahuano
- 9f*** †Taino
- 9g** †Caquetio
- 9h** †Shebayo

10 North Amazonian*Colombian*

- | | |
|----------------------------|----------------------|
| 10a †Resigaro | 10d Piapoco |
| 10b Yucuna (†Guarú) | 10e Cabiayari |
| 10c Achagua | 10f †Maipure |

Upper Rio Negro

- 10g** Baniwa of Içana/Kurripako
- 10h** Tariana
- 10i** Guarequena

Orinoco

- 10j** Bare
- 10k** Baniwa of Guainia
- 10l** †Yavitero (Baniwa of Yavita)
- 10m** †Mandawaka
- 10n** †Yabaana

Middle Rio Negro

- 10o** Kaixana
- 10p** †Manao
- 10r** †Bahwana/Chiriana

† = extinct, * = outside of the scope of the map

I have been able to consult data for a bit more than a half of the languages, viz. some 27 (for a couple of languages, more than one variety) of the 38 living languages and 5 of the 22 extinct languages. For some 24 of these languages, both from the North Arawakan and South and Southwestern Arawakan groups, the source consulted report the possibility of a general ‘who?, what?’ interrogative. The languages with the presumed ‘who?, what?’ interrogatives are summarized in Table 1.

Table 1. Arawakan languages reported to have a ‘who?, what?’ interrogative

	(*)	Forms & Sources
Terêna	1a	<i>kuti</i> (Maria de Lourdes Sobrinho via Onilda Sanches Nincao, p.c.; Eastlack 1968:7; GICLI Swadesh list for Terêna)
Baure	1e	[((AG-)woyi)ko(wo)n + usually a deictic (typically, the gender-number neutral <i>to</i> ART) + sometimes <i>ka</i> REL], the most common realization being <i>kon to</i> (Swintha Danielsen, p.c.)
Waurá	2a	<i>kaitsa</i> (GICLI Swadesh list for Waurá)
Mehinaku	2b	<i>atsa</i> (GICLI Swadesh list for Mehinaku)
Yawalapiti	2c	<i>kanauka</i> (GICLI Swadesh list for Yawalapiti)
Maniteneri/Mantinera	3a	<i>tëhirë</i> (Ramirez 2001:614) ⁱ⁾
Apurinã	3c	<i>ki-pa</i> , <i>ke-pa</i> , <i>ke-AG-pa</i> (da Silva Facundes 2000:123, 364-366, 541-542; cf. Section III.8.2.3)
Pichis Asheninca, Apurucayali Asheninca ⁱⁱ⁾	4b	<i>tsika(ri-ka)</i> (also ‘where?’; alternative spellings: <i>ç^hika</i> and <i>t^{sh}ika</i>), <i>(tsika)</i> AG- <i>paita(rika)</i> /AG- <i>paita(ri)ka</i> /AG- <i>itarika</i> , ⁱⁱⁱ⁾ <i>paita(ka)</i> (Payne <i>et al.</i> 1982:230, 240; Reed & Payne 1986:328-329; for references and examples, cf. also Cysouw, forthcoming)
Ucayali-Yurúa Asheninca	4b	<i>himpe/hempe</i> ^{iv)} (also ‘where?’; Payne 1980:71; Reed & Payne 1986:330)
Nomatsiguenga	4e	<i>paí-AG</i> , viz. <i>paí-rí</i> M and <i>paí-ró</i> F (Shaver 1996:37, 169; cf. Section III.8.2.3)
Wapixana/Wapishana	7a	[<i>kanum</i> (+ <i>qí</i> : DEM)], [<i>kan</i> + <i>qí</i> : DEM] (Gomes dos Santos 2006:193-194, 234-235)
Island Carib/Iñeri (dialect: Dominica)	9a	<i>kát(a)</i> - (see Section III.7.6)

Garifuna/Black Carib/ Central 9b	<i>ka(-), káta-</i> (see Section III.7.6)
American Carib (dialects: Hopkins in Belize and Livingston in Guatemala)	
Arawak/Lokono	9c <i>(h)ama</i> (de Goeje 1928:32, 177) ^{v)}
Yucuna	10b <i>na</i> (Schauer & Schauer 2000:526, 532; Ramirez 2001:579)
Achagua	10c <i>tana</i> (Meléndez Lozano 2000:631, 639; Wilson & Levinsohn 1992:26, 125-127), <i>tanáa</i> (Ramirez 2001:324)
Piapoco	10d <i>ka(w)iná</i> (also ‘how?’; Ramirez 2001:296, 572), <i>kaináa</i> (also ‘how?’; Mosonyi 2000:655-656)
Cabiyari/Cabiyari/Kabiyari	10e <i>mikhá</i> (Ramirez 2001:389)
Maipure	10f <i>iti</i> (Zamponi 2003:52)
Baniwa of Içana	10g <i>kúaka/k^huaka</i> (in some contexts may be just <i>kúa</i> ; Ramirez 2001:163, 195, 221, 568; Taylor 1993:151, 167; GICLI Swadesh list for Baniwa)
Guarekena/Warekena	10i <i>ifi</i> (Aikhenvald 1998:261, 325-326)
Baré/Bare	10j <i>ne</i> (Aikhenvald 1995b:25; GICLI Swadesh list for Baré)
Mandahuaca/Mandawaka	10m <i>kaniaka</i> (Ramirez 2001:572)
Kaishana/Kaixana	10o <i>napa-hi</i> (Nimuendajú 1941 via Ramirez 2001:432-433) ^{vi)}
(dialect: Rio Tonantins)	

* The number-letter combinations indicate the classification and the location of the language in the same way as is done for Map 2.

- i) There may be dialectal differences, since other sources give distinct forms: *kli* ‘what?’ vs. *katni* ‘who?’ (GICLI Swadesh list for Machineri), *klu-he* ‘what?’ vs. *kat-he* ‘who?’, *kat-ni-he* ‘who.PL?’ (Maia *et al.* 2000:300 for Manxineri), *katu* ‘who?’ vs. *klu* ‘what?’ (Wise 1986:573 for Piro).
- ii) The Perené dialect of Asheninca distinguish *ninka* ‘who?’, *paíta* ‘what?’, and *tsika* ‘where?’ (Reed & Payne 1986:329-330).
- iii) In principle, the last form without *pa* should be available only in the Apurucayali dialect, because only in the Apurucayali (and the Ucayali-Yurúa) dialect, the verb ‘name, call’, used in this construction, has the root *-ii(t)-*, while in Pichis Asheninca it has the form *-pai(t)-* (where *t* is epenthetic; Payne 1980:62, 99; Payne *et al.* 1982:117-124, 246).
- iv) The possibility of its use as ‘who?’ is suggested by the sources only indirectly and, therefore, needs to be checked.
- v) According to Marie-France Patte (p.c.; cf. also Patte 2002), in the modern language the interrogative *hama(a)* means only ‘what?’.
- vi) Another source cited by Ramirez (2001:432-433) for the Kaishana variety that used to be spoken at Lake Mapari (lower reaches of Rio Japurá), Hanke (1960), gives distinct forms: *hökama* ‘what?’ vs. *nakamáhi* ‘who?’.

By way of illustration, examples (1-3) demonstrate the use of such interrogatives for Baure, Apurinã and Baniwa of Içana.

Baure (non-North Arawakan; Bolivia; Swintha Danielsen, p.c.)

- (1) a. *ro-woyikowon?*
3SG.M-IPW
'Who/what is it?'
- b. *woyikon = iš ri-kotoriač-on-ow to ni-šonon [kač ri-hakak]?*
IPW=well 3SG.F-play.with-NMLZ-COP ART 1SG-daughter.in.law
'Who/what is my daughter-in-law playing with [that she is laughing]?'
- c. *woyikon teč vi-nik-pa?*
IPW DEM.M.SG 1PL-eat-GO
'What are we going to eat?'
- d. *kon to pi-nik?*
IPW ART 2SG-eat
'What will you eat?'
- e. *kon to wono-vi?*
IPW ART send-2SG
'Who sent you?'

Apurinã (non-North Arawakan; Brazil)

- (2) a. *ke-(ru)-pa po-ka-pe?*
IPW-(3M)-Q 2SG-kill-PFV
'What/who have you killed?' (da Silva Facundes 2000:365, 542)
- b. *ke-ru-pa umaka?*
IPW-3M-Q sleep
'Who sleeps?' (da Silva Facundes 2000:365)
- c. *u-kara iye ke-ru-pa? [pūtanuru pitximunanhi]*
3M-DISTAL then IPW-3M-Q
'[A:] What is it?! [B: 'The penis body of your husband!']'⁶ (da Silva Facundes 2000:491)
- d. *ki-pa apo-pe?*
IPW-Q arrive-PFV
'Who/what has arrived?' (da Silva Facundes 2000:365)

⁶ This question "is uttered after one of the woman's family member[s], in disapproval of the woman's bond with the tapir, cuts off the tapir's genitalia and throws it at the woman" (da Silva Facundes 2000:491).

Baniwa of Içana (North Arawakan; Brazil, Colombia, Venezuela)

- (3) a. *kúa-ka í-uka-Ji-wa?*
 IPW-SUB 3-come-REL-MIDDLE.VOICE
 ‘Who is coming? (lit.: ‘Who is it who is coming?’)’ (Ramirez 2001:195)
- b. *kúa-ka pi-kaíte-Ji?*
 IPW-SUB 2SG-say-REL
 ‘What do you say? (lit.: ‘What is it that you say?’)’ (Ramirez 2001:221)

As can be readily observed in Table 1, the lack of differentiation between ‘who?’ and ‘what?’ appears to be very common among the Arawakan languages. However, what is even more remarkable is that the forms of the ‘who?, what?’ interrogatives vary greatly from language to language. In fact, this is true not only for the ‘who?, what?’ interrogatives, but also for other interrogatives. Thus, in her overview of the Arawakan languages, Aikhenvald (1999:85) notes that “interrogatives vary even among closely related languages”. Similarly, discussing the pronominals of Asheninca, Reed & Payne (1986:328) remark that although “a paucity of forms is [...] characteristic of the [Asheninca] interrogative and indefinite pronouns [...] this is probably one of the areas of greatest dialect divergence in Asheninca”.

A detailed comparative study would be necessary to account for the details of the variation in the Arawakan interrogatives and to trace the origins of both the lack of differentiation between ‘who?’ and ‘what?’ in some languages and its presence in the others. However, on the whole, this variation appears, similarly to the Mayan languages discussed in Section III.7.5, to be mostly due to the tendency to build the interrogative pronominals as complex constructions, which are later truncated in various ways (Section III.8.2.1).⁷ Furthermore, in some cases, the variation can be explained by the development of the locative interrogatives ‘where?’ into selective and subsequently non-selective interrogative pronominals (Section III.8.2.2). Moreover, the peculiarities of the semantics of genders and agreement patterns may have played a role in some cases (Section III.8.2.3). It is not possible to discuss all this in detail here and only a brief illustration of the mechanisms mentioned will be provided in what follows.

8.2.1 Complex constructions and their truncation

Let us begin by considering the tendency of the Arawakan languages to build

⁷ A possible implication of this is that even if any originally simplex interrogatives did exist in Proto Arawakan, they might have been lost without traces in most, if not all modern languages.

their interrogatives as constructions (and subsequently truncate them) on the example of Asheninca and some other non-North Arawakan languages.

Asheninca (Pichis and Apurucayali dialects) has the general interrogative *tsika* (also spelled as *ɛ^hika*, *t^{sh}ika*) ‘where?, who?, what?, which one?, how?, why?’. According to Reed & Payne (1986:329), its “unmarked semantic interpretation [...] seems to be ‘where?’” (this claim is also supported by Cysouw, forthcoming), as in (4).

(Pichis?) Asheninca (non-North Arawakan; Peru)⁸

(4) *tsika mula-payeeni?*

IPW mule-PL

‘Where are the mules?’ (Anderson 1985-1986:I.136, reference via Cysouw, forthcoming)

Most other meanings of *tsika* can be explained as derived from its primary locative meaning ‘where?’. Semantically, this would not be implausible. Thus, the non-selective ‘who?’ and ‘what?’ (and from ‘what?’ to ‘why?’) can be derived from ‘where?’ via the selective ‘which one?’. The use of ‘where?’ as ‘how?’ is also plausible, both typologically and semantically. What is special about Asheninca, though, is that these semantic extensions coexist simultaneously for one form and that unlike most other languages, Asheninca does not have any other underived interrogative prowords. Yet, in Asheninca, as in most other languages, it is perfectly possible to specify the intended interrogative meaning. Asheninca does this by combining *tsika* with a small number of verbs, which provide more explicit cues for its exact interpretation (cf. Cysouw, forthcoming for a discussion and references). Thus, to express the interrogative pronominal meanings ‘who?, what?, which one?’, *tsika* is usually combined with the verb ‘name, call’, *-pai(t)-* (most typical for the Pichis dialect) or *-ii(t)-* (most typical for the Apurucayali dialect), so that the whole construction literally means something like ‘what/who does it call itself that [P]’, (5).

(Pichis?) Asheninca

(5) a. *tsika o-pai-t-a-ka h-antz-i-ri?*

IPW 3F-name-EP-NON<FUT>.REFL-Q 3M-do-NON<FUT>-REL

‘What is he doing? (lit.: ‘What does it call itself that he does?’) (Reed & Payne 1986:329)

b. *tsika o-pai-t-a-ka antz-i-ro-ri?*

IPW 3F-name-EP-NON<FUT>.REFL-Q do-NON<FUT>-3F-REL

‘Who did it?’ (Reed & Payne 1986:329)

⁸ I have slightly modified the spelling of the Asheninca examples to make them uniform.

- c. *tsika i-pai-t-a-ka pok-atz-i-ri*
 IPW 3M-name-EP-NON<FUT>.REFL-Q come-STATIVE-NON<FUT>-REL
iroñaaka?
 now
 ‘Who is coming?’ (Anderson 1985-1986:I.136, reference via Cysouw, forthcoming)

The combination of *tsika* and the verb ‘name, call’ is to a considerable degree conventionalized and *tsika* is regularly dropped, as in (6a), sometimes together with the gender-number marker, as in (6b) (cf. also Section III.8.2.3).

(Pichis?) Asheninca

- (6) a. *i-pai-t-a-ka p-antz-i-ri-ka?*
 3M-name-EP-NON<FUT>.REFL-Q 2SG-do-NON<FUT>-REL-Q
 ‘What are you doing?’ (Anderson 1985-1986:II.42, reference via Cysouw, forthcoming)
- b. *pai-t-a-ka p-antz-i-ri-ka?*
 name-EP-NON<FUT>.REFL-Q 2SG-do-NON<FUT>-REL-Q
 ‘What are you doing?’ (Anderson 1985-1986:II.110, reference via Cysouw, forthcoming)

In the Perené dialect of Asheninca, this construction has resulted in the interrogative *paita* ‘what?’ (Reed & Payne 1986:330), as in (7).

Perené Asheninca (Reed & Payne 1986:330)

- (7) *paita p-antz-i-ri?*
 what 2SG-do-NON<FUT>-REL
 ‘What are you doing?’

The Nomatsiguenga ‘who?, what?’ interrogatives *paí-rí* (M)/*paí-ró* (F) (cf. Table 1) must have a similar origin.

Cysouw (forthcoming) analyzes *tsika*, as a combination of the interrogative root *tsi-* and the interrogative suffix *-ka*. He also suggests that “the Asheninca interrogative root *tsi-* might be related to the suffix *-(n)tsi*, which indicates indefinite possessor”, viz. it is “used with inalienable nouns to indicate that the possessor is not known or not of importance”. However, I do not find this hypothesis convincing. At least, not in this form, because the mechanisms that might account for the evolution from the nominal indefinite possessor suffix to the general interrogative (with ‘where?’ as the primary meaning) remain unexplained.

Instead, it may be argued that just like *paita(ka)* and similar interrogatives, the interrogative *tsika* itself may also go back to a verb, thus pointing to a recurrent pattern in the diachronic development of Arawakan interrogatives. This

possibility is suggested by the fact that, for instance in Apurucayali Asheninca, *tsika* can be marked with *-ri-ka* -REL-Q, as in (8a), just as the verb *-ii(t)-* ‘name, call’ in (8b), where the latter is used in the same way as *-pai(t)-* ‘name, call’ in (6a) above.

Apurucayali Asheninca

- (8) a. *tsika-ri-ka h-impoi-t-ant-apiint-a-ri*
 IPW-REL-Q 3M-follow-EP-BECAUSE-HAB-NON<FUT>.REFL-REL
apaniroidini?
 alone
 ‘Why does he always go behind alone? (lit.: ‘What is it for that he always follows alone?’)’ (Payne *et al.* 1982:240)
- b. *o-i-t-a-ri-ka h-ant-i-ri?*
 3F-name-EP-NON<FUT>.REFL-REL-Q 3M-do-FUT-REL
 ‘What will he make?’ (Payne *et al.* 1982:230)

Among the verbs that *tsika* may go back to, the most plausible candidate seems to be the verb *-vitsik-* [-βit^hik-] (Apurucayali dialect)/-(*o*)*vetsik-* (Ucayali-Yurúa dialect) ‘make’ (Payne 1980:97; Payne *et al.* 1982:36, 230, 236). The vowel *-a* of *tsik-a* would then be the same NON<FUT>.REFL marker *-a-* as in *o-i-t-a-ri-ka* in (8b).

The origin of *tsika* in the verb ‘make’ is not incompatible with the fact that the unmarked semantic interpretation of *tsika* appears to be the locative ‘where?’. In this respect, compare the use of the combination [-*k(i)e-* ‘say, do, happen’ + *-po* PFV.REFL] in the meaning ‘where?’ in Baure (Swintha Danielsen, p.c.), as in (9).

Baure (Swintha Danielsen, p.c.)

- (9) a. *no-kie-po-n to pohi-nev?*
 3PL-**do**-PFV.REFL-NMLZ ART duck-PL
 ‘Where are the ducks?’
- b. *no-kie-po-wo-n to pohi-nev?*
 3PL-**do**-PFV.REFL-COP-NMLZ ART duck-PL
 ‘Where are the ducks going?’

Verbs meaning both ‘say’ and ‘do’ (sometimes, also ‘be’) are common in Amazonian languages in general and in Arawakan languages in particular (Swintha Danielsen, p.c.; cf. Payne 1990:79). Furthermore, semantically, the meanings ‘do’ and ‘make’ are quite close as well. In this respect, compare, for instance, the Baure examples (10a), where the verb *-kič-* ‘say, do’ is used, and (10b), where the verb *-woyik-* ‘make’ is used instead in exactly the same meaning.

Baure (Swintha Danielsen, p.c.)

- (10) a. *hare' ni-piri, pi-kič-ow-on?*
 hello 1SG-sibling 2SG-**do**-COP-NMLZ
 'Hello, sister, what are you doing?'
 b. *ri-woyik-on = iš?*
 3SG.F-**make**-NMLZ=well
 'What is she doing then?'

The hypothesis that Asheninca *tsika* goes back to the verb 'make' may be further supported by the fact that the verb 'make' also appears to be used to form interrogative pronominal constructions in at least one other non-North Arawakan language, viz. Terêna. Thus, according to Eastlack (1968:7), in Terêna, "sometimes, instead of *kuti* ['who?, what?'], we find the interrogative phrase *kuti itukóvo*, which doubtless has its origin in a subject emphasis transformation of a sentence of the type *itukóvo šoáum* 'John made himself, John became, John is...'; thus its [literal] meaning would be 'who is the one who became...' [h]owever it is found to substitute for noun phrases and to mean 'which one?'" as in (11).

Terêna (non-North Arawakan; Brazil; Eastlack 1968:7)

- (11) *kuti itukó-vo koépeko šoáum?*
 IPW made-REFL killed PROP
 'Which one did John kill?'

Yet another non-North Arawakan language, where the form expressing the interrogative meanings 'who?' and 'what?' has been clearly built as a verbal construction, is Baure. (Some verbal constructions with other interrogative meanings have already been mentioned for Baure above, (9-10)). In Baure, according to Swintha Danielsen (p.c.), the meanings 'who?, what?, which one?, which [N]?' are expressed by the construction [((AG-)*woyi*)*ko(wo)n* + usually a deictic (typically, the gender-number neutral *to* ART) + sometimes *ka* REL], which most frequently is realized as *kon to* (cf. examples (1) above). However, due to some features of the Baure morphonology it is difficult to say with certainty which verb has served as the basis for this construction. What is clear, though, is that the final *-n* is a nominalizer and *-wo* is a "copula"⁹ (Swintha Danielsen, p.c.). According to Swintha Danielsen (p.c.), there are at least three more or less equally plausible possibilities. First, the base may be constituted by the verb *-woy-* 'be called, be named' and the so-called "absolute" (or "absolutive") "verb base suffix" *-ko* (Swintha Danielsen, p.c.; verbal "stem formative",

⁹ This morpheme, going back to the locative verb *-wo-* 'be (somewhere)', has various functions. For instance, Baptista & Wallin (1967:66) label *-wo* "punctiliar" and "verbalizer".

Baptista & Wallin 1967:66).¹⁰ Second, the base may be the verb *-woyik(o)-* ‘make’, which itself “seems to have the root *-woyi-* plus the absolutive suffix *-ko*”. Finally, the third possibility is that the base is made up of the locative verb *-wo-* ‘be’ and the verbal locative suffix *-(i)yi/- (i)y(o)*,¹¹ which is used for “locative subordination”, as in (12a), or for asking the question ‘where?’, as in (12b, c). In the latter case, the verb must be subsequently nominalized.

Baure

- (12) a. *noy teč in n-er-íy-ow no-ro-pík-ów*
 there DEM.M.SG water 3PL-drink-LOC-COP 3PL-just.now-come-COP
to neč ka?áno-neb
 ART DEM.PL animal-PL
 ‘Those animals were just now going there where they drink that water’
 (Baptista & Wallin 1967:32).
- b. *ro-komoroki-yi-vi-n kove?*
 3SG.M-bite-LOC-2SG-NMLZ dog
 ‘Where did the dog bite you?’ (Swintha Danielsen, p.c.)
- c. *r-avi-y-wo-n to pi-wer?*
 3SG.M-live-LOC-COP-NMLZ ART 2SG-house
 ‘Where is your house?’ (Swintha Danielsen, p.c.)

In the preceding discussion, quite a few examples have already been provided to illustrate the tendency in Arawakan languages for the formal reduction of originally more complex interrogatives. By way of conclusion, two more apparent cases of such a reduction from Table 1 may be mentioned. Thus, compare Waurá *kaitsa* ‘who?, what?’ vs. Mehinaku *atsa* ‘who?, what?’ and also Achagua *tana* ‘who?, what?’ vs. Yucuna *na*. Waurá and Mehinaku are closely

¹⁰ The terms *absolute* and *absolutive* used to refer to *-ko* do not have anything to do with the expression of case. According to Swintha Danielsen (p.c.), “it is the morpheme that you can derive independent nouns with, as e.g. *ni-wohis* (1SG-hand) ‘my hand’ and *wohiso-k(o)* (hand-ABSOLUTE) ‘hand or hand-like thing’. The same morpheme is used as a verb base suffix. It is related to possible objects that are related to the action, which is not identical with transitivity”. Baptista & Wallin (1967:66) use the term *absolute* for *-ko* without any further explanations, but they put it in the same category of the verbal “stem formatives” as the “instrumental” (applicative) *-čō* and the “indefinite, pseudo passive” *-sí*.

¹¹ The basic variant of this suffix is *-yi*, which may sometimes undergo metathesis, epenthetic vowel insertion and change of *i* to *e* (when there is an *o* in the preceding or the following syllable). The verbal locative suffix is “phonologically slightly different from the locative marker on NPs, which is *-ye*, but it seems obvious that both forms go back to the same source” (Swintha Danielsen, p.c.).

related and even mutually intelligible,¹² the relation between Achagua and Yucuna is somewhat less close.

8.2.2 ‘Where?’, ‘which one?’, ‘who?’, and ‘what?’

At least in some Arawakan languages, the non-selective interrogative pronominals ‘who?’ and ‘what?’ appear to result from the earlier locative interrogatives ‘where?’, in all probability through the selective meaning ‘which one?’. One such possible example has already been pointed out in Section III.8.2.1, viz. the Asheninca interrogative *tsika*, which has the locative ‘where?’ as its “unmarked semantic interpretation” (Reed & Payne 1986:329), but can also be used as ‘who?, what?, which one?, how?, why?’. Another example may be provided by the Lokono/Arawak interrogative *hali-ka*-AG ‘who?, which one? (person or thing)’ (Northern Arawak, Caribbean; Suriname; Marie-France Patte, p.c.; Patte 2002:92), which is clearly related to one of the two Garifuna ‘where?’ interrogatives, viz. *halia*- (Northern Arawak, Caribbean; Belize, Guatemala, Honduras, Nicaragua; Taylor 1956a:23).¹³ In Achagua, the interrogative *tana* ‘who?, what?, which one?’ seems to be related to the interrogative *ta/tanakutfa* ‘where?’ (cf. Meléndez Lozano 2000:631-632, 635, 639).¹⁴ In Machiguenga, the animate interrogative *tya-ni* ‘where is he/she/it (AN)?, which one (AN)?’ (lit.: be.where?-AN) is also used as the non-selective human interrogative ‘who?’,¹⁵ instead of the dedicated ‘who?’ interrogative *tsini*, which is nowadays used only infrequently (Betty Snell, p.c.).¹⁶

8.2.3 Gender

As can be observed in Table 1, provided in the beginning of Section III.8.2, in

¹² “The Waujá variant of Aruak [i.e., the Arawakan language Waurá] is intelligible to speakers of Mehinako [= Mehinaku]” (cf. Thomas Gregor, 2002, at www.socioambiental.org/pib/epienglish/mehinako/lang.shtm).

¹³ The other Garifuna ‘where?’ interrogative is *hagá*- ‘where?’ (Taylor 1956b:139). The two interrogatives differ primarily in their morphosyntactic properties.

¹⁴ The interrogative *tanakutfa* ‘where?’ may contain the nominal suffix *-ku* ‘in’ (Meléndez Lozano 2000:632-633).

¹⁵ Its inanimate counterpart *tya-ti* ‘where is it (INAN)?’ can, it seems, be used only selectively as ‘which one (INAN)?’ (Betty Snell, p.c.). The locative interrogative ‘where?’ unmarked for animacy is *tya-ra* (Snell 1998:70-71). The final *-ra* here seems to be the locative nominalizer, as in *no-mag-i-ra* (1SG-sleep-REAL-LOC.NMLZ) ‘where I sleep, my bed’ or *i-nori-a-ra* ‘where he is lying down’ (3M-lie.down-REAL.REFL-LOC.NMLZ) (Snell 1998:68-69).

¹⁶ In fact, the interrogative *tsini* appears to be used so rarely that it has not even been included in Snell’s (1998) dictionary of Machiguenga.

some languages the ‘who?, what?’ interrogatives may, or need to be marked for gender-number (and person), viz. Baure, Apurinã, Asheninca, and Nomatsiguenga. Most Arawakan languages have gender.¹⁷ What is more, quite a few of them have a so-called combined gender system (Corbett 1991:184). That is, in such languages different agreement features may control different sets of agreement patterns, usually on different targets (cf. Van de Velde 2006:204-205). In Arawakan languages, the divide is typically between “pronominal modifiers and verbal cross-referencing” on the one hand, and “the remaining modifiers”, on the other (Aikhenvald 2000:70). In the first case, which pertains more directly to the interrogative pronominals, the distinction is normally made between *masculine* and *feminine* genders. It should be pointed out, however, that one of the two genders is usually functionally unmarked, so that the opposition is rather either between masculine vs. non-masculine, as in Lokono (cf. Patte 2002:91), or feminine vs. non-feminine, as in Guarekena/Warekena (Aikhenvald 1998:297). Typically, gender assignment is determined by such parameters as sex, humanness and animacy. For instance, in Apurucayali Asheninca, male humans and most animates are masculine, while female humans and most inanimates are feminine (Payne *et al.* 1982:48). Agreement patterns may sometimes be used for purposes other than the expression of gender agreement. Thus, in Lokono (Patte 2002:91), the masculine gender contains nominals referring to male humans, while all other nominals, including those referring to women, are assigned to the general non-masculine gender. However, the “masculine marker [...] is also used to mark any highly individuated entity, including feminine ones”.

Although the sources are rarely explicit on this point, it is to be expected that the semantics of genders and agreement patterns may interact with the semantics of the interrogatives used to convey the meanings ‘who?’, ‘what?’ and ‘which one?’. In what follows, I will briefly consider this issue on the example of Lokono, Nomatsiguenga, Apurinã and Asheninca.

In Lokono, whose gender system has been just described, the interrogative *hali-ka-AG* ‘who?, which one? (person or thing)’ “generally occurs with the non-masculine marker *-n*, namely as *halikan*” (“although its masculine counterpart *halikai* is also attested”; Patte 2002:93), even when the speaker is perfectly aware that the person in question is male (Marie-France Patte, p.c.). In this respect, consider (13a) and (13b).

¹⁷ For an overview of gender and other means of nominal classification in Arawakan languages see, for instance, Aikhenvald (1994, 1999:83-84).

Lokono/Arawak (Northern Arawak, Caribbean; Suriname; Marie-France Patte, p.c.)

- (13) a. *halika-n to-ra?*
 IPW-NON<M> NON<M>-DEM
 ‘Who is that (person)?’
 b. *halika-n li-ra?*
 IPW-NON<M> M-DEM
 ‘Who is that (*man*)?’

Nomatsiguenga, as described in Shaver (1996), has a combined gender system, distinguishing masculine vs. feminine on the one hand, and animate vs. inanimate, on the other (Shaver 1996:29-30). However, only the former opposition is (obligatorily) marked with the interrogative pronominals, viz. *paírí* M vs. *paíró* F (Shaver 1996:37). Humans are distributed between the masculine and feminine genders according to their sex, animates may belong to both genders,¹⁸ and most inanimates are reported to be assigned to the feminine gender (Shaver 1996:29). Judging from this description, one would expect the masculine form of the interrogative, *paírí*, to be used in questions about humans (default and males) and animals (default and males) and the feminine form, *paíró*, to be used in questions about women, female animals, things, etc. Shaver (1996) does not discuss this issue and translates both the masculine and the feminine forms as ‘who?, what?, which?’ (“qué?, quién?, cuál?”, 1996:37). However, the examples he provides, such as (14), suggest that *paírí* and *paíró* in Nomatsiguenga may in fact be functionally differentiated along the lines just hypothesized.

Nomatsiguenga (non-North Arawakan; Peru)

- (14) a. *paí-rí ha-t-atsí?*
 IPW-M go-EP-STATIVE.NON<FUT>
 ‘Who is it going (there)?’ (Shaver 1996:37)¹⁹
 b. *paí-ró Ø-óg-ë-mi?*
 IPW-F 3F-take-REAL-2SG
 ‘What have you fallen ill with?’ (lit.: ‘What (a sickness) has taken you?’) (Shaver 1996:169)

¹⁸ Although Shaver (1996) does not discuss this point explicitly, the masculine seems to be the default option for non-human animates, somewhat like in the closely related language Asheninca (cf. above in the present section).

¹⁹ Unlike in the following example, no cross-reference prefix, not even a zero one, is glossed on the verb here because the “stative” (“estativo”) suffix is incompatible with prefixal cross-referencing on the verb (cf. Shaver 1996:60).

c. *paí-ró pi-hi-t-a?*

IPW-F 2SG-call-EP-REAL.REFL

‘What is your name?’ (lit.: ‘What do you call yourself?’) (Shaver 1996:37)

Apurinã, as described by da Silva Facundes (2000), appears to distinguish feminine vs. non-feminine (“masculine”) genders.²⁰ The non-selective interrogative pronominal ‘who?, what?’ can (i) be left unmarked for gender, viz. *ki-pa/ke-pa*, (ii) be marked for the masculine gender, *ke-ru-pa*, or the feminine gender, *ke-ro-pa* (da Silva Facundes 2000:123, 364-366, 541-542). Examples of the form unmarked for gender and the masculine form have already been provided in (2). For convenience sake, I reproduce them here as (15).

Apurinã (non-North Arawakan; Brazil)

(15) a. *ke-(ru)-pa po-ka-pe?*

IPW-(3M)-Q 2SG-kill-PFV

‘What/who have you killed?’ (da Silva Facundes 2000:365, 542)

b. *ke-ru-pa umaka?*

IPW-3M-Q sleep

‘Who sleeps?’ (da Silva Facundes 2000:365)

c. *u-kara iye ke-ru-pa?* [*pūtanuru pitximunanhi*]

3M-DISTAL then IPW-3M-Q

‘[A:] What is it?! [B: ‘The penis body of your husband!]]’ (da Silva Facundes 2000:491)

d. *ki-pa apo-pe?*

IPW-Q arrive-PFV

‘Who/what has arrived?’ (da Silva Facundes 2000:365)

The form unmarked for gender, which means ‘who?, what?’, is not problematic. Equally, the masculine form *ke-ru-pa* appears to behave as would be expected on the basis of the semantics of the masculine gender in Apurinã, viz. it appears to be used in questions about humans (default and males), animals (default and males), and things. The feminine form would be expected to be used in questions about women, and possibly, (clearly) female animals. Unfortunately, da Silva Facundes does not say anything on this issue. However, the only example of the feminine form *ke-ro-pa* I found in da Silva Facundes (2000:123, also repeated on

²⁰ Thus, consider the following statements made by da Silva Facundes (2000:224-225): “the great majority of [...] stems referring to sex-non-differentiable elements is assigned to the masculine gender”; “the **default gender** [bold in the original] is the masculine one. The default gender is the one imposed on noun forms when these lack any morphological gender marking, semantic or pragmatic indications of their gender”.

page 628) is translated as ‘Who is she?’, which is in conformity with the expectations.

Finally, let us consider the gender marking in the Asheninca interrogative constructions *(tsika) i-paitaka* (M) and *(tsika) o-paitaka* (F), which were discussed in Section III.8.2.1. Cysouw (forthcoming), who has analyzed 206 constituent questions extracted from the collection of the Asheninca texts by Anderson (1985-1986), reports to “have not found any reason for the choice between these two alternatives [i.e., the masculine *i-* and the feminine *o-*]”. According to Cysouw, “there appears to be a free choice between the use of either prefix”. However, as far as I can judge from Cysouw’s sample, provided at the end of his (forthcoming) paper, the choice between the masculine *i-* and the feminine *o-* is not completely random. Admittedly, it is not completely predictable either, but certain tendencies are positively discernable. As expected, on the whole, these tendencies appear to go along the same lines as the gender assignment tendencies mentioned earlier in this section for (Apurucayali) Asheninca. Recall that male humans and most animates are treated as masculine and female humans and most inanimates as feminine (Payne *et al.* 1982:48).

The functional differentiation between the masculine *i-* and the feminine *o-* in the interrogative construction *(tsika) AG-paitaka* appears to be most consistent in the examples where *tsika* is present, which are covered by types A.6.1-A.6.5 in Cysouw’s sample. In particular, the construction with the masculine marker, *tsika i-paitaka*, appears to be primarily used in questions about humans (default and males)^{21, 22} and animals (one example involving a bird and three examples explicitly or implicitly referring to animals as meat), while the construction with the feminine marker, *tsika o-paitaka*, appears to be used in questions about

²¹ It should be mentioned that I analyze as questions about the identity of persons not only the questions subsumed under Cysouw’s type A.6.2 “*tsika* + *-pait-* + relative clause (‘who’)”, translated as ‘Who is coming?’ and the like, but also those questions subsumed under Cysouw’s type A.6.1 “*tsika* + *-pait-* (‘call how’)” that inquire about the name of a person, even though in the source they are translated as ‘How is he called?’ and the like. The reason is that in both cases the interrogative constructions used are in principle the same [*tsika* + *-pait-*]. In A.6.2, the basic construction [*tsika* + *-pait-*] is just further elaborated by means of a dependent predication. In other words, just as the examples in A.6.2 can be translated as something like ‘How is he called the one who is coming?’ instead of the original ‘Who is coming?’, the examples in A.6.1 translated originally as ‘How is he called?’ can be reinterpreted as ‘Who is he?’.

²² In the latest version of Cysouw’s paper available to me (15 February 2005), the last example under A.6.1 *tsika ipaitaka noñaakeri* ‘How is he called the one I saw?, Who was it I saw?’ (Anderson 1985-1986:III.86) has been misprinted with the feminine marker *o-*, viz. *tsika opaitaka noñaakeri*.

women (one example) and inanimate entities or concepts (the remaining examples).

It is only in the examples where *tsika* is omitted, which are covered by types A.7.1-A.7.5 and A.8 in Cysouw's sample, that the suggested distribution becomes distorted. The only exception here is type A.7.1 “-*pait*- + relative clause ('who')”,²³ where again, as when *tsika* would have been present (type A.6.2), the masculine marker *i*- is used in four of the five examples. What is more, in the only example in which *i*- is not used, *o*- is not used either, instead the bare form *paitaka* is found. It is probably hardly a coincidence that when *tsika* is omitted, the bare form *paitaka* unmarked for gender (and/or person) becomes quite common as well. There are 21 such bare forms in the examples when *tsika* is absent and none in the examples where *tsika* is present. Finally, only in the examples where *tsika* is present, any non-third person marking, such as *pi*- 2SG, occurs in Cysouw's sample.

8.3 Chibchan languages: Arhuaco/Ika

The languages of the Chibchan family are spoken both in North and South America stretching from Nicaragua in the north to the border regions of Colombia and Ecuador in the south. Since “some of the most fundamental diversity internal to the family is found in Costa Rica and western Panama” and “the closest presumable relatives of the Chibchan family as a whole, Lenca and Misumpala, are located at the northern [...] borders of the Chibchan domain”, Chibchan is likely to be of North American origin (Adelaar 2004:50). The internal classification of the Chibchan languages found in the Ethnologue is rather flat, with nine equidistant branches and one unclassified language.

One Chibchan language has been found to lack differentiation between ‘who?’ and ‘what?’. The language at issue is primarily known in the literature as Arhuaco or Ika, the latter name being used by the people themselves. Ika, as described by Frank (1985) and Landaburu (2000), is spoken on the southern slopes of Sierra Nevada de Santa Marta in the north of Colombia and belongs to the so-called Aruak/Arhuac(an) branch of the Chibchan family. The other two languages of this branch, Cogui/Kogui and Damana/Malayo/Arosario/Wiwa/Guamaca/Sanja/Sancá, are spoken immediately to the north of Ika.

The Ika interrogative ‘who?, what?’ is reported by Frank (1985:113-114) as *ini* and by Landaburu (2000:747) as *inə*, but the difference between the two forms is only a matter of spelling conventions. According to Frank (1985:113-114), *ini* ‘who?, what?’ “may refer to a subject, object or oblique NP” and it may

²³ That is, AG-*paitaka* (without the preceding *tsika*) when it is used to convey the meaning ‘who?’.

be used with “the suffix/clitic *-aba*’, which is also found attached to words for days of the week, borrowed from Spanish”, the resulting combination being *iniba* ‘what day (of the week)?, when?’. The use of the interrogative *ini* ‘who?, what?’ is illustrated with the examples in (16).

Arhuaco/Ika

- (16) a. *ini was-i-ri ei ž-ăn no?*
 IPW chase-while-TOP thus say-IPFV Q.NON<PST>
 ‘What is it chasing, barking like that?’ (Frank 1985:113)
- b. *ini sin nă-nas-e?*
 IPW with 2SG-come-Q.PST
 ‘Who did you come with?’

8.4 Makú languages: Dâw

The so-called Makú/Maku languages are spoken by compact groups of nomadic hunter-gatherers in the border region between Brazil and Colombia. According to Martins & Martins (1999:251), the Makú family consists of “four languages belonging to seven tribes”, whereas the Ethnologue distinguishes six Makú languages.

One Makú language has been found to allow a lack of differentiation between ‘who?’ and ‘what?’. The language at issue is Dâw, as described by Martins (2004). Dâw is spoken by a group of almost a hundred persons on the right bank of the Upper Rio Negro in the vicinity of the town of São Gabriel. According to Martins (2004:375, 384-387), the Dâw interrogative *paj/pàj* may mean both ‘what?’ (17a) and ‘who?’ (17b), as well as ‘why?’ (17c).²⁴

Dâw

- (17) a. *pàj tuk ?ãm fěh?*
 IPW want 2SG nephew
 ‘What do you want, nephew?’ (Martins 2004:375)
- b. *pàj ?ãm ?xj ka??*
 IPW 2SG call PROG
 ‘Whom are you calling?’ (Martins 2004:386)
- c. *pàj ?ãm ?ýt?*
 IPW 2SG weep
 ‘Why are you crying?’ (Martins 2004:386)

²⁴ Martins (2004:386): “os diferentes significados que ele [i.e., *pàj*] pode expressar: tais como *quem?*; *o quê?*; *por quê?*” (“the different meanings that it [i.e., *pàj*] may express, such as *who?*, *what?*, *why?*”).

There is also a dedicated human interrogative *huĩ?* ‘who?’ (Martins 2004:384-385, 387-388), as in (18).

Dâw

- (18) *huĩ? ?a ?e?*
 IPW DEM PTCL
 ‘Who is that, I wonder?’ (Martins 2004:387)

Martins (2004:385-386) suggests that *paj/pâj* ‘may be bimorphemic’ in origin, viz. *p-âj*, where *-âj* may be the same as the locative suffix *-âj* ‘more in front, a bit distant’, as in (19).

Dâw

- (19) *?a-hõt-âj t-êh puíd fug kafãm jet*
 DEM-far-more.in.front far-NEG be.EMPH PROP die be.lying.down
xrd-ãm
 DURATIVE-TELIC
 ‘A bit further ahead, Xigui was going to die’ (Martins 2004:386)

8.5 Urarina

Urarina, as described by Olawsky (2006), is a linguistic isolate spoken by less than 3000 people in the area of the Río Chambira, in the department of Loreto in northern Peru. According to Olawsky (2006:258-259, 814-831), Urarina uses only one interrogative pronominal *dʒa* ‘who?, what?, what (kind of) [N]?, which [N]?’ as illustrated in (20).

Urarina

- (20) a. *dʒa kwara-i tahia?*
 IPW see-2 over.there
 ‘Whom/what have you seen over there?’ (Olawsky 2006:816)
- b. *dʒa tũũ-e?*
 IPW arrive-3_E
 ‘Who has come?’ or possibly ‘What has arrived?’ (Olawsky 2006:815)
- c. *dʒa kurete-i na-ĩ baha-anũ?*
 IPW buy-2 say-PTCP ask-1SG_A
 ‘“What have you bought?”, I asked him’ (Olawsky 2006:816)
- d. *dʒa=na itça-kũre nii rihiheĩ nii ka=iri-ũũ?*
 IPW-Q do-3PL_E that like that 1SG=creature-PL
 ‘What [the heck] have these creatures of mine done?’ (Olawsky 2006:816)

e. *dʒa itɕa-re = i?*

IPW do-IRR.3_E=ASSERTIVE

‘What would he do?’ or ‘Who would do it?’ (Olawsky 2006:817)

Most other interrogatives are based on *dʒa*, such as *dʒa raj* ‘whose?, for whom?’ (*raj* ‘POSS, for’), *dʒa-rihi* ‘which one?’ (*rihi* ‘be like’), *dʒa-nu* ‘why?’ (*-nu* PURPOSIVE), *dʒa-toane-i* ‘how?’ (*-i* PTCP; *ka-toania* ‘be like this’, *ni-toania* ‘be like that’), *dʒa-bana* ‘when?’ (*bana* ‘at the time when’), *dʒ-u* ‘where?’ (*-u* LOC) (Olawsky 2006:258-259, 814-825). The form *dʒa* may also function as an indefinite non-human pronoun ‘something’ (in negative clauses, ‘nothing’; Olawsky 2006: 827-828). Remarkably, it cannot be used in the human meaning ‘someone’. Instead, ‘someone’ (in negative clauses, ‘nobody’) is based on the locative form of *dʒ-u* (*<*dʒa-u* IPW-LOC) as *dʒ-u-ne(j)/dʒ-u-ni*, where *-ne(j)/-ni* is “only found in this context” (Olawsky 2006: 828-829).

It may be hypothesized that the interrogative *dʒa* ‘who?, what?’ is related to the verb *itɕa* ‘do’, or more precisely to its inflected form *itɕa-a* ‘do-3_A’. To begin with, such a verbal origin involving verbs like ‘do, make’ would not be exceptional for the region. In this respect, recall some non-North Arawakan languages discussed in Section III.8.2.1. Second, the alternation between [dʒ] and [tɕ] is “non-phonemic [, since they] occur in complementary distribution” (Olawsky 2006:31). The voiced affricate realization [dʒ] “occurs word-initially, before /a/ and /u/”, while the voiceless alveopalatal affricate realization [tɕ] “mainly occurs word-internally, as a syllable onset” (Olawsky 2006:39). Furthermore, the vowel /i/ in “the sound sequence /i/ + /tɕ/ [...] is mute in rapid speech” (Olawsky 2006:40) and “the prototypical syllable type found in Urarina is CV” (Olawsky 2006:75). Finally, the semantics of the “person inflectional class” suffix *-a* 3_A would also fit here. The so-called “person inflectional class” markers in Urarina are verbal inflectional suffixes expressing person and number of the argument cross-referenced on the predicate, as well as the reality status, polarity, focus, and the degree of finiteness of the predicate (cf. Olawsky 2006:431-454). There are three such person inflectional classes “A-form”, “E-form” and “D-form”. The “A-form”, such as *-a* 3_A, is obligatory among other things in the “citation form” of the verb, “in greetings”, and importantly, “in polar questions” (Olawsky 2006:432, 441-445). What is more, it is “overwhelmingly typical” in “‘short’ utterances”, which “can be defined as [utterances] that consist[...] of an inflected verb only or an inflected verb with maximally one NP or adverbial” (Olawsky 2006:443).

8.6 Arauan languages: Paumari

The Arauan (or Arawá) language family consists of six languages spoken in the south of the state of Amazonas, Brazil, in neighbouring parts of the state of Acre

and just over the border in Peru. There is one Arauan language, Paumari, as described by Chapman (1986, 1988) and Chapman & Derbyshire (1991:203-216), that has a single invariable interrogative pronominal *nahina* ‘who?, what?’ (21), which can also be used attributively as ‘what [N]?, which [N]?’ (22).

Paumari (Chapman 1988)

- (21) a. *nahina* *'ida* *kana-ni-ra* *nofi-ja?*
 IPW DEM.F.SG bathe-NMLZ.F.SG-OBJ want-INDEP
 ‘Who wants to take a bath?’
- b. *nahina-ra* *'i-no'a-vini* *hi-ja* *'ida* *ihai?*
 IPW-OBJ 2SG-give-NMLZ be-INDEP DEM.F.SG medicine
 ‘Whom did you give the medicine?’
- c. *nahina* *mani* *'ida* *Maria* *'i-ra* *no'a-vini* *hi-ja?*
 IPW COP DEM.F.SG PROP 2SG-OBJ give-NMLZ be-INDEP
 ‘What did Maria give you?’
- d. *nahina* *mani* *'oni?*
 IPW COP DEM.F.SG
 ‘What is this (near you)?’
- e. *nahina* *ka-so'oro* *mani* *'oni?*
 IPW 3SG.GEN-basket COP DEM.F.SG
 ‘Whose is this basket?’ or ‘What is this basket for?’
- (22) a. *nahina* *gora* *'i-okha-ki-i?*
 IPW house 2SG-go-BACKGROUND-2SG
 ‘To which house are you going?’
- b. *nahina* *mahi* *bana* *'i-okha-ki-li?*²⁵
 IPW day future 2SG-go-BACKGROUND-2SG
 ‘What day will you go?’

According to Salzer & Chapman (1996), the same form *nahina* can also be used as an indefinite pronominal, both human ‘somebody’ and non-human ‘something’, as well as in the meaning ‘thing’ (cf. also Chapman 1988) and ‘(be) unconscious, (be) in coma’.

The interrogative *nahina* is invariable in gender-number. In plural contexts, the plural form *hana-vi-hi-na* ‘who? (PL), what? (PL), which ones?’ (lit.: where-3PL-be-NON<3SG.F>) of the otherwise selective interrogative ‘which (one)?’ must be used (Chapman 1988). In the singular, the selective interrogative has the forms *hana-hi-ni* ‘which (one)? (3SG.F)’ and *hana-hi-na* ‘which (one)? (NON<3SG.F>, which in the absence of the plural marker *vi-* implies M.SG)’ (Chapman 1988). Note that *nahina* ‘who?, what?’ is very similar to the non-

²⁵ [sic!] Shouldn’t the form be *'i-okha-ki-i?*

feminine singular form of the selective interrogative *hanahina*.

8.7 Chapacura-Wanham languages: Itene/More

Traditionally, the languages of the Chapacura-Wanham family are spoken in the department of Beni in northern Bolivia and in Brazil in the adjacent state of Rondônia and further north in the south of the state of Amazonas. Most known Chapacura-Wanham languages are moribund or extinct. The Ethnologue reports five presumably still extant Chapacura-Wanham languages, which it subdivides into two branches, the Guaporé river branch and the Madeira river branch.

One Chapacura-Wanham language Itene/More has been found to allow a lack of differentiation between ‘who?’ and ‘what?’. Itene, as described by Angenot-de Lima (2002), is a moribund language traditionally spoken in Bolivia in the region to the south of the confluence of the Guaporé and Mamoré rivers. According to Angenot-de Lima (2002:282, 334-337, 458, 733), Itene uses only one general interrogative pronominal *?ati(:)/?atin(:)/?etin*, which corresponds to the underlying forms $|?a: = tin|$ (2002:282, 334-337, 458) or $|?a: = ti:|$ (2002:458). This interrogative may on itself mean ‘who?’ (23a), ‘what?’ (23b), ‘what [N]?, which [N]?’ (24), ‘why?’ (25), and in combination with some further material, ‘when?’ (26), ‘how many?, how much?’ (27), and ‘(do) how?’ (28) (also ‘(say) what?, how?’).

Itene²⁶

- (23) a. *?ati:* *tənˀ* *nā(nˀ)* *kˀewakˀ?*
 $|?a: = tin\#$ $tən\#$ $na:-?a\#$ $kawak|$
 IPW# make# IPFV-N# pirogue[N]
 ‘Who made the pirogue?’ (Angenot-de Lima 2002:335)
- b. *?ati:* *tənˀ* *tənˀ* *kˀa?*
 $|?a: = tin\#$ $tən\#$ $tən\#$ $ka:|$
 IPW# make# make# 3NON(N)
 ‘What are you doing/making?’ (the original French translation: ‘Que

²⁶ The first line in these examples reflects the surface realization (“réalisation phonétique”), while the second line provides the presumed underlying form (“représentation (morpho)phonologique”, Angenot-de Lima 2002:15). The juncture markers in the underlying form have the following values: ‘#’ word-external boundary (“limite externe de mot”), ‘≠’ strong word-internal boundary (“limite interne forte de mot”), ‘=’ weak word-internal boundary (“limite interne faible de mot”), and ‘-’ affixation boundary (“limite d’affixe”; Angenot-de Lima 2002:15). I added the gender of the nominals between [square brackets] in the gloss line. Literal translations have also been added by me.

fais-tu?’)²⁷ (Angenot-de Lima 2002:335)

- (24) *ʔati* **itən*^ʔ *nɔk^xa?*
 |ʔa: = ti: # ʔi: = tən # na:-ʔɔn = ka:|
 IPW# person# SUB-M=3NON<N>
 ‘Who is he? (lit.: ‘What person (is) he?’)’ (Angenot-de Lima 2002:458)
- (25) *ʔati:* *waw^{ʔw}* *waw^{ʔw}* *ma?*
 |ʔa: = tin # waw^{ʔw} # waw^{ʔw} # ʔum = na:|
 IPW# cry# cry# 2=IPFV
 ‘Why are you crying?’ (Angenot-de Lima 2002:335)
- (26) *ʔatin*^ʔ *nɔm*^ʔ *najə* *tut*^ʔ *ma?*
 |ʔa: = tin # ʔɔtɔm # na:-ʔaɲ = ji:-ʔip # tus # ʔum = na:|
 IPW# FUT# SUB-N=3N-N# walk# 2=IPFV
 ‘Till **when** are you going to walk? (lit.: ‘**When** will it have become that you are walking?’)’ (Angenot-de Lima 2002:336)
- (27) *ʔatin*^ʔ *nɔk^xɔ* *pati* *pa:* *jə* *sɛ:mɛ?*
 |ʔa: = tin # na:-ʔɔn = ka:-ʔɔn # pa: = tiʔ # pa: # ji:-ʔip # sɛ: ≠ mɛ:|
 IPW# SUB-M=3NON<N>-M# fish[M] # kill # 3N-N# caiman[N]
 ‘How many fish did the caiman kill? (lit.: ‘How many has he [i.e., the fish] become that the caiman killed?’)’ (Angenot-de Lima 2002:337)²⁸
- (28) *ʔatin*^ʔ *najə* *tən*^ʔ *nɔk^xaɲ*^ʔ *k^xewak*^ʔ?
 |ʔa: = tin # na:-ʔaɲ = ji:-ʔip # tən # na:-ʔɔn = ka:-ʔɔn-ʔaɲ # kawak|
 IPW# SUB-N=3N-N# make# SUB-M=3NON<N>-M-N pirogue[N]
 ‘How did he make the pirogue? (lit.: ‘How has it become that he made the pirogue?’)’ (Angenot-de Lima 2002:337, 309)

The only interrogative meaning that seems to be expressed without *ʔati(:)/ʔatin(:)/ʔetin* appears to be the locative ‘where?’ (29). Note that in this case no other interrogative root is used either. Rather, the meaning ‘where?’ is expressed by the construction as a whole (presumably, with the respective interrogative intonation).

²⁷ Note that the 3NON<N> pronominal *k^xa* seems to suggest the translation ‘What is *he* doing?’ rather than ‘What are *you* doing?’.

²⁸ I have changed the gloss of |na:-ʔɔn = ka:-ʔɔn #| from IPFV-M=3NON<N>-M#, which appears to be a typo, to SUB-M=3NON<N>-M#, in accordance with the analysis by Angenot-de Lima (2002:309) and by analogy with the gloss in a similar example that follows this example in the source.

Itene

- (29) *ʔi:* *ɔma: *jə* *ɸutup*⁷ *jə* *m^wɛjak*⁷?
 |ʔi:p# ʔɔma:# ji:-ʔip# pu: = tup# ji:-ʔip# m^wijak⁷|
 N# be# 3N-N# go.out# 3N-N# peccary[N]
 ‘Where did the peccary go out? (lit.: ‘There is it (the place) that the peccary went out?’)’ (Angenot-de Lima 2002:337-338)

With its paucity of interrogatives, Itene resembles some non-North Arawakan languages, also spoken in the southwestern fringes of the Amazonia, as discussed in Section III.8.2.1. Expectedly, Itene is also quite similar in this regard to other Chapacura-Wanham languages, such as Wari’/Pakaásnovos/Pakaas Novos. Wari’, as described by Everett & Kern (1997:18-33, 42-52), does not have dedicated interrogatives at all. Instead, Everett & Kern (1997) argue that Wari’ uses (non-proximate) demonstratives in combination with “inflectional morphemes” (i.e. gender/tense/(ir)realis markers) in the so-called “complementizer construction”, as in (30).

Wari’

- (30) a. *ma’* *co* *querec* *ma*?
 that.near.you NON<N>.REAL.NON<FUT> see 2SG
 ‘Whom did you see?’ (Everett & Kern 1997:25)
- b. *ma’* *ca* *querec* *ca*?
 that.near.you N.REAL.NON<FUT> see 3SG.M
 ‘What did he see?’ (Everett & Kern 1997:46)
- c. *ma-in* *capam’* *ca* *xo’* *’ep* *ma*?
 that.near.you-3N cornbread[N] N.REAL.NON<FUT> recently grind 2SG
 ‘Which cornbread did you just make?’ (Everett & Kern 1997:50)
- d. *ma-in* *ca* *mao* *ca*?
 that.near.you-3N N.REAL.NON<FUT> go.SG 3SG.M
 ‘Where did he go?’ (Everett & Kern 1997:50)

The differentiation between the human and non-human meanings of the interrogative construction appears to be primarily achieved by the gender marker used, non-neuter (i.e., feminine or masculine) markers imply the human reading and the neuter marker the non-human reading. Note the similarity between the construction meaning ‘which [N]?’ in (30c) and the construction meaning ‘where?’ in (30d).

It may also be interesting to point out that the form *ma’* used in (30) is not only a demonstrative but also an existential predicate ‘there is, there exists’, as in (31), apparently derived from the demonstrative.²⁹

²⁹ Besides *ma’*, in constituent questions Wari’ may use *cain’* to “express the notion of ‘what?’

Wari'

- (31) *ma' na xin wari' co*
 EXIST 3SG.REAL.NON<FUT> 1PL.INCL>3N person NON<N>.REAL.NON<FUT>
ho' pawin non quem ne?
 put.in height 3SG.REAL.NON<FUT>>3SG.M PTCL RECENT.PST
 'Somebody has made a hiding place to lie in wait for him, hasn't he? (lit.:
 'There exists a person who has made a hiding place...')' (Everett & Kern
 1997:16)

What is more, the “inflectional morphemes” *co* and *ca* appear to function as a kind of relative pronouns in both (30) and (31). All this suggests that *ma'* in (30) should probably be better interpreted as an existential predicate, so that for instance in (30a, b), the whole construction can be interpreted as something like ‘There is ... who/that P?’ with the conventionalized implicature ‘Who/what is he/she/it who/that P?’. Such an existential interpretation of *ma'* in constituent question constructions would also make the Wari' locative interrogative construction in (30d) structurally much more similar to the respective Itene locative interrogative construction in (29) above.

The important role of gender markers in the interrogative constructions of Wari' may suggest an explanation for the lack of differentiation between ‘who?’ and ‘what?’ in Itene. Thus, in Itene, there is a strong resemblance between one of the two underlying variants of the neuter marker *|-?ijɲ|* and one of the three underlying variants of the feminine marker *|-?in|* (Angenot-de Lima 2002:277).³⁰ The latter form is also (almost) identical with the neuter marker *-in* in Wari', as in (30c-d). Furthermore, in Wari', the feminine is also used for “collective nouns” and “mixed groups of female and male individuals” (Aikhenvald & Dixon 1999:360). Finally, both in Wari' (30d) and in Itene (29) the neuter is used in the locative interrogative construction meaning ‘where?’. All this seems to suggest, that (i) either the final segment in the Itene interrogative *?ati(:)/?atin(:)/?etin* may be due to a conflation of the feminine and neuter markers or (ii) that it is an originally neuter marker and that *?ati(:)/?atin(:)/?etin* used to be a locative interrogative ‘where?’ (similar to the Wari' *ma-in ca* (30d)),

(as applied to a proposition, e.g. the content of an utterance or thought), ‘when?’, ‘how?’, ‘how many?’ and ‘how long?’, as well as to question the predicate, i.e. as ‘do what?, be how?’ (Everett & Kern 1997:40, 51). Note that *cain'* is a “verb zero-derived from a demonstrative (*cain'* ‘to be that (N.DISTAL)’)” (Everett & Kern 1997:40) and in this respect it appears to be similar to *ma'*.

³⁰ The other underlying forms are *|-?aɲ|* for the neuter and *|-?an, -?am|* for the feminine (Angenot-de Lima 2002:277). The variants with the vowel *i* seem to be used when the preceding syllable has an *i* as well.

which was later extended to cover most other interrogative meanings³¹ and subsequently replaced by a new construction (29) in its original locative interrogative meaning ‘where?’.

8.8 Tacanan languages

The five living languages of the Tacanan family are spoken in northwestern Bolivia and the adjacent regions of Peru. According to the Ethnologue, the Tacanan family consists of two branches, the Araona-Tacana branch including Araona, Cavineña, Reyesano and Tacana, and the Ese Ejja branch represented by a single language. I have been able to check data for all the five living languages. The languages found to use a single interrogative ‘who?, what?’ are summarized in Table 2.

The use of ‘who?, what?’ interrogatives in Tacanan languages may be illustrated with the example of Cavineña *ai* (32).

Cavineña

- (32) a. *ai = tsewe = tu e-tata = ke = ra shana-kware e-bakwa = ke?*
 IPW=ASS=3SG 3-father=3=ERG leave-REMOTE.PST 3-child=3
 ‘[A:] Who did the (lit.: ‘his_i’) father leave his child_i with? [B: He left him with his (the child’s) grandfather’.]’ (Guillaume 2004:103)
- b. *tume ai = tsewe yu-wa kemi-ya? wika = tsewe?*
 then IPW=ASS over.there-LOC take.out-IPFV hook=ASS
mailla = tsewe? tarafa = tsewe? u ai = tsewe?
 net=ASS casting.net=ASS or IPW=ASS
 ‘What do you fish (lit.: ‘take fish out’) with over there (in your country)? With hooks? With a net? With a casting net? Or with what?’ (Guillaume 2004:533)
- c. *ai = jatsu = tu-ja = tu ani-ya?*
 IPW=EXACTLY=3SG-DAT=3SG sit-IPFV
 ‘What is it exactly that he (the sick person) has? (lit.: ‘What exactly sits to him?’)’ (Guillaume 2004:445)
- d. *ai = mi ju-ya Kana?*
 IPW=2SG be-IPFV PROP
 ‘[When the Cavineña saw Kana, the traitor, back near the Cavineña village, he asked him:] What are you doing Kana?’ (Guillaume 2004:104)

³¹ In this respect, compare the Asheninca (non-North Arawakan; Peru) interrogative *tsika* discussed in Section III.8.2.1.

Table 2. Tacanan languages reported to have a ‘who?, what?’ interrogative

Forms & Sources	
Cavineña	<i>ai</i> ‘who?, what?, what (kind of) [N]?’ with = <i>kwana</i> PL as ‘things’ ⁱ⁾ with = <i>kwana</i> DUB as an indefinite ‘something, somebody’ or as a filler (placeholder) ‘whatsi(t)sname’ (Guillaume 2004:81, 102-105, 358, 444-445, 533, 537, 543, 562, etc.)
Tacana	<i>ai</i> ‘who?, what?’ ⁱⁱ⁾ <i>a(‘)ya</i> ‘who?’ <i>hai</i> ‘what?’ (Ottaviano & Ottaviano 1967:150, 175-176, 179, 201)
Ese Ejja (Baawaja Kuinañe dialect, Peru) ⁱⁱⁱ⁾	<i>‘a’ia/a-</i> ‘who?, what?’ (María C. Chavarria Mendoza, p.c.)
Ese Ejja (Bolivia)	<i>ae/a-</i> ‘who?, what?, what [N]?, which [N]?’ <i>aiya/aide</i> ‘who?’ (Shoemaker & Shoemaker 1967:223, 228, 257, 276)

i) In Guillaume (2004), the root *ai* in the meaning ‘thing’ appears to be found only with the plural marker =*kwana*.

ii) It seems that most commonly it is used just as ‘what?’.

iii) Spoken on the Baawaja/Tambopata river in Peru.

Cavineña

- (32) e. *ai jae=mi kemi-wa?*
IPW fish=2SG take.out-PRF
‘What type of fish did you catch?’ (Guillaume 2004:445)

The remaining two Tacanan languages, Reyesano and Araona, do not allow a lack of differentiation between ‘who?’ and ‘what?’. Thus, Reyesano has *aise* ‘who?’ and *ai* ‘what?’ (Antoine Guillaume, p.c.) and Araona has *aise/aiyaha* ‘who?’ (also ‘somebody’) and *ai* ‘what?’ (also ‘something’ and ‘thing’) (Pitman 1981:22, 24).

It seems that the lack of differentiation between ‘who?’ and ‘what?’ should be reconstructed for Proto Tacanan, the form of the respective interrogative in the proto language being in all probability closest to that of the Peruvian Ese Ejja interrogative *‘a’ia* ‘who?, what?’. Dedicated ‘who?’ interrogatives, such as

Tacana *aiya/aide* ‘who?’,³² developed later on the basis of such a general ‘who?, what?’ interrogative. This development must have resulted in the subsequent specification of the original ‘who?, what?’ interrogative to the non-human function ‘what?’, as in Tacana *ai*, which mostly seems to mean ‘what?’, but sometimes still can be used as ‘who?’ (e.g., *ai sa* ‘whose? (lit.: IPW GEN)’ Ottaviano & Ottaviano 1967:175-176).

As far as the origin of the Proto Tacanan ‘who?, what?’ interrogative itself is concerned, at least two hypotheses are possible. First, as suggested by such forms as Cavineña *ai* ‘who?, what?, what (kind of) [N]?’ and *ai=kwana* ‘things’ or ‘something, somebody, whatsi(t)sname’, this ‘who?, what?’ interrogative may have originated as the word ‘thing’ used as a filler (placeholder) ‘whatsi(t)sname?’ both for things and persons, comparably to French *machin*, Flemish *dinge(s)* or Arabic *šayʿ*² (cf. Section III.2.3.4.2.1). Later it has been conventionalized as an interrogative pronominal ‘who?, what?’. Another fact that can be viewed as supportive of such a non-interrogative origin of the Tacanan ‘who?, what?’ interrogatives is that for instance in Cavineña, “interrogative clauses [...] do not appear to have any specific interrogative intonation nor any obligatory question marker that would mark these clauses as different from basic clauses” (Guillaume 2004:102).

Alternatively, it may be hypothesized that, as is not unattested in the region (cf. Section III.8.2.1 on some non-North Arawakan languages, Section III.8.5 on Urarina), this interrogative may be verbal in origin. In particular, it may be related to the widespread Tacanan verbal root *a*, such as Araona *a* ‘do, make, say’ (Pitman 1981:11), the Cavineña “auxiliary” *a* ‘affect’ (Guillaume 2004:65, 123-124, 148), Tacana *a* ‘do’ (Ottaviano & Ottaviano 1967:192, 198), or Ese Ejja *a* ‘do, say’ (Shoemaker & Shoemaker 1967:250). The final part in the Tacanan ‘who?, what?’ interrogatives, such as Ese Ejja *’a’ia*, may then be some tense-aspect marker, comparable for instance to the Cavineña imperfective marker *-ya*. The non-interrogative meanings of the Tacanan forms also used as ‘who?, what?’ interrogatives may then be due to later semantic expansions.

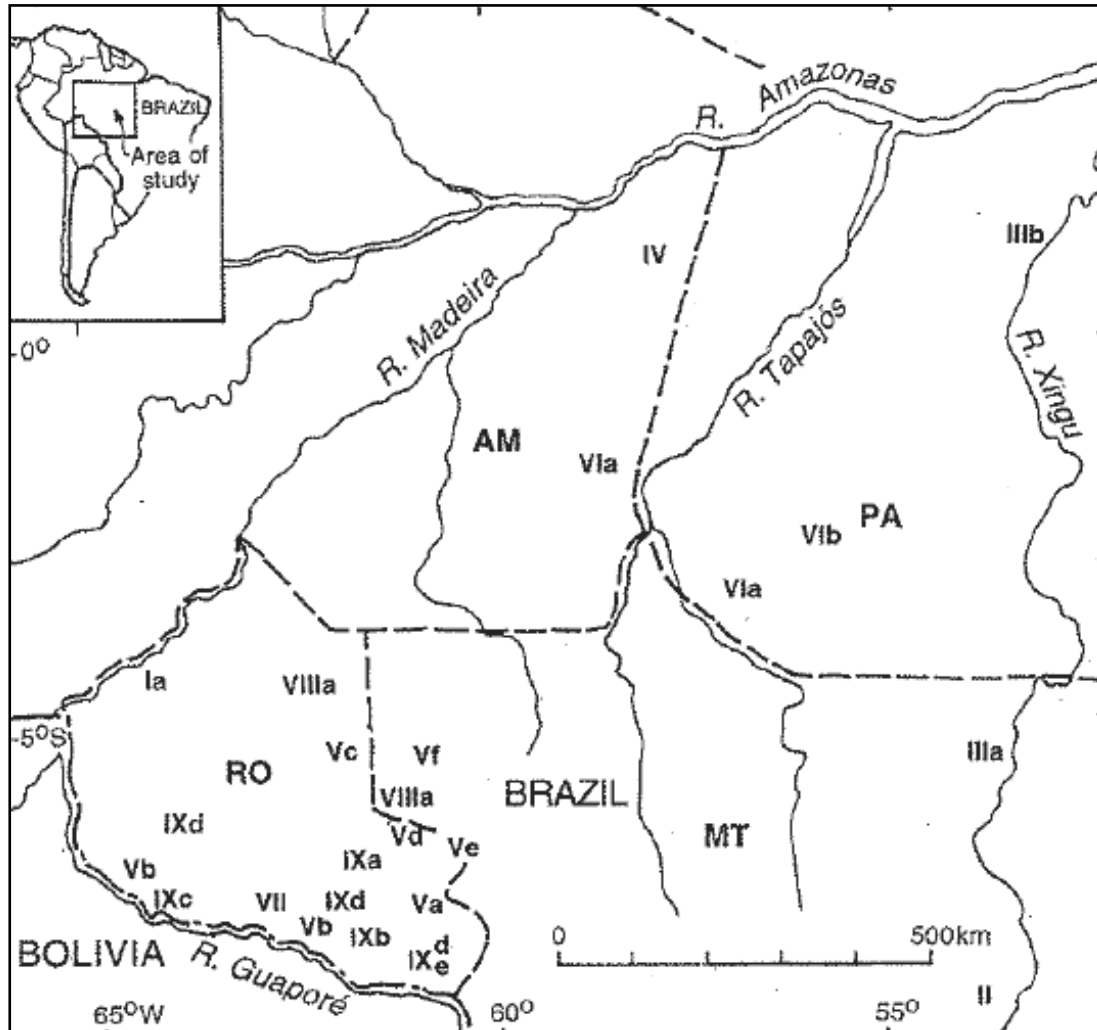
8.9 Tupí languages

The Tupí family is one of the biggest linguistic groups in South America. According to Rodrigues (1999a:107-110), it consists of ten branches, the largest and the best-described being the Tupí-Guaraní branch. Tupí-Guaraní languages are spoken throughout Brazil (with the exception of the northeastern parts), as well as in Paraguay, Argentina, Bolivia, Peru, and French Guiana. The remaining

³² For instance, Tacana *aide* ‘who?’ appears to result from the contraction of **ai deha* ‘what/which man?’.

nine branches have just a few members each and are all concentrated in western Brazil, primarily in the state of Rondônia and the neighbouring areas to the northeast and east of it, as illustrated on the Map 3.³³

*Map 3. (Extant) non-Tupí-Guaraní languages of the Tupí family with approximate locations (following Rodrigues 1999a:108-109)**



³³ The spellings of the language names used in this section may differ slightly from those used by Rodrigues (1999a) and reproduced here in the legend to the Map 3.

I Arikém branch

- a Karitiána
- b †Arikém

II Awetí branch

- a Awetí/Aueti/Auetö

III Juruna branch

- a Juruna/Yuruna/Yudya
- b Xipáya
- c †Manitsawá

IV Mawé branch

- a Mawé/Maué/Sateré-Mawé

V Mondé branch

- a Mondé
- b Aruá
- c Gavião/Ikõro/Digüt
- d Suruí/Paitér
- e Cinta-larga
- f Zoró

VI Mundurukú branch

- a Mundurukú
- b Kuruáya

VII Puruborá branch

- a Puruborá

VIII Ramaráma branch

- a Káro
- b †Urumí

IX Tuparí branch

- a Tuparí
- b Wayoró/Ajurú
- c Mekéns
- d Makurap
- e Sakirabiát
- f †Kepkiriwát

X Tupí-Guaraní branch**

† = extinct

* Two capital letter abbreviations on the map refer to the Brazilian states, viz. *RO* stands for Rondônia, *MT* for Mato Grosso, *AM* for Amazonas and *PA* for Pará.

** Not marked on the present map.

I have been able to consult some data for at least one language from each Tupí branch, except the Puruborá and Ramaráma branches. Table 3 summarizes the four languages from three branches that appear to allow a lack of differentiation between ‘who?’ and ‘what?’. The ‘who?, what?’ interrogatives of Karitiána (Arikém branch), Gavião and Suruí (both Mondé branch) in Table 3 appear to be cognate with each other, but not with the Sakirabiá ‘who?, what?’ interrogative *arop*. The link between the first three forms is further supported by the existence of forms such as Cinta Larga *mê* ‘who?’ and *mênã* ‘what?’ (Tupí, Mondé; GICLI Swadesh list for Cinta-Larga). As to Sakirabiá *arop* ‘who?, what?’, in the closely related language Mekéns the same form *arop* means ‘what?’, with ‘who?’ being expressed by *apo* (Galucio 2001:166-168). The other two languages of the Tuparí branch I have data for have the following forms: (i) Tuparí has *apo* ‘who?’ and *kat?at* ‘what?’ (GICLI Swadesh list for Tuparí), and (ii) Makuráp has *arikop* ‘what?’ (GICLI Swadesh list for Makuráp; no form for ‘who?’ is provided).

Table 3. Tupí languages reported to have a ‘who?, what?’ interrogative

	(*)	Forms & Sources
Karitiâna/Karitiána	Ia	<i>môra</i> ‘who?, what?’ <i>môn(ã)</i> ‘what?, what [N]?, which [N]?’ <i>mônramôn</i> ‘what?’ (Landin 1984:15-16, 2005:16, 48, 51; Landin 1987:201)
Gavião (do Jiparaná/ do Rondônia)/Ikôro/ Digüt	Vc	<i>mé</i> ‘who?, what?’ (GICLI Swadesh list for Gavião do Jiparaná)
Suruí/Paitér	Vd	<i>nan</i> ‘who?, what?’ <i>kaná</i> ‘what?’ <i>kána</i> ‘where?, what?’ <i>káná</i> ‘why?’ <i>ká</i> ‘which?’ (http://geocities.com/indianlanguages_2000/tupi.htm#Monde – a wordlist compiled by Victor A. Petrucci)
Sakirabiá(t)	IXe	<i>arop</i> ‘who?, what?’, also ‘thing’ (Galucio 2001:166-168)

* The number-letter combinations indicate the classification and the location of the language in the same way as is done for Map 3.

8.10 Macro-Jê languages

The languages of the so-called Macro-Jê (or Macro-Ge) stock are all spoken in Brazil. “[S]everal [...] languages are spoken in Brazilian Amazonia, [but] the geographical distribution of this linguistic stock is rather circum-Amazonian, encircling Amazonia on its eastern and southern sides” (Rodrigues 1999b:166). The approximate locations of the extant Macro-Jê languages are provided on Map 4 (see below).³⁴ It should be pointed out that, according to Rodrigues (1999b:165), Macro-Jê is not a fully established genetic entity, rather it is “a [commonly used] working hypothesis whose details have varied according to different scholars”. Thus, the Ethnologue’s classification of Macro-Jê includes three linguistic groups not included by Rodrigues (1999b), viz. (i) Chiquitano/Chiquito family, consisting of a single language spoken in Bolivia,

³⁴ The spellings of the language names used in this section may differ slightly from those used by Rodrigues (1999b) and reproduced here in the legend to the Map 4.

(ii) Yabutí family, consisting of two languages, Arikapú and Jabutí, both spoken in the state of Rondônia, Brazil, and (iii) the extinct single-language Oti family, which used to be spoken in the state of São Paulo, Brazil. In the present section, I will follow Rodrigues's (1999b) classification of the "possible members of Macro-Jê", which is reproduced here as the legend to Map 4.

I have been able to consult data for all the eight Macro-Jê families with living members. At least one variety of each language, except Panará/Kreen-Akarore/Kren-akarôre and Umutína, has been considered. Four of these seventeen languages appear to allow a lack of differentiation between 'who?' and 'what?': Xerente/Xerênte (Central Jê; Section III.8.10.1), Krenak/Krenák/Botocudo/Borun (Krenák; Section III.8.10.2), Maxakalí (Maxakalí; Section III.8.10.3) and Fulniô/Yatê (Yatê; Section III.8.10.4). Of these four, Xerente provides the most solid case. In what follows, I will present the data of these Macro-Jê languages in more detail.

8.10.1 Xerente

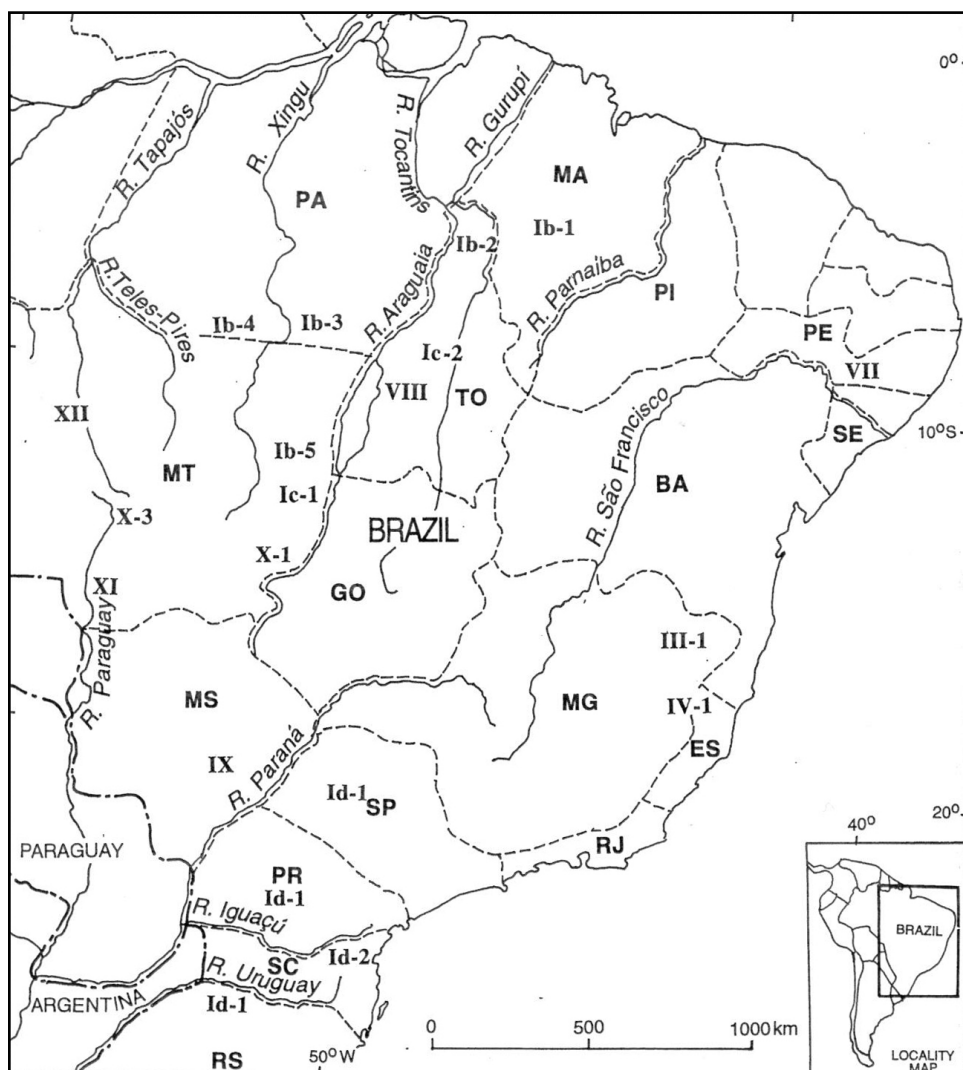
Xerente/Xerênte (or Akwê-Xerente) is a central Jê language spoken by around two thousand people in the central part of the state of Tocantins, Brazil. My data on Xerente interrogatives come from Sinval Martins de Sousa Filho (p.c.).

There exist two sex-based varieties of Xerente, the so-called men's speech and women's speech. The differences between the two are not very big. However, Xerente appears to be typologically unique in that 'who?' and 'what?' are distinguished in men's speech, viz. *nôkwa* 'who?' (33) vs. *mãrĩ* 'what?' (34), but not in women's speech, where a single form *taha* 'who?, what?' is used (35).

Xerente (men's speech; Sinval Martins de Sousa Filho p.c.)

- (33) a. *nôkwa b-za wara?*
 who Q-FUT run
 'Who will run?'
 b. *nôkwa mã-p ai-kwakri?*
 who PST-Q 2-prick
 'Who pricked you?'
 c. *nôkwa-p to-ta ambi?*
 who-Q QUOTATIVE-DEM man
 'Who is that man?'

Map 4. (Extant) Macro-Jê languages with approximate locations (following Rodrigues 1999b:164, 167-168)*



I Jê family**a Northeastern Jê**

- 1 †Jaikó

b Northern Jê

- 1 Timbira (Canela, Krahô, Gavião of Pará/Par(a)katêjê, Kri(n)katí(-Timbira))
- 2 Apinajé
- 3 Kayapó
- 4 Panará/Kreen-Akarore/Kren-akarôre
- 5 Suyá

c Central Jê

- 1 Xavánte
- 2 Xerénte
- 3 (†)Xakriabá
- 4 †Akroá

d Southern Jê

- 1 Kaingáng
- 2 Xokleng
- 3 †Ingaín

II Kamakã family

- 1 †Kamakã
- 2 †Mongoyó
- 3 †Menién
- 4 †Kotoxó
- 5 †Masakará

III Maxakalí family

- 1 Maxakalí
- 2 †Kapoxó
- 3 †Monoxó
- 4 †Makoni

- 5 †Malalí

- 6 †Pataxó

IV Krenák family

- 1 Krenák/Botocudo
- 2 †Guerén

V Purí family

- 1 †Purí
- 2 †Koropó
- 3 †Coroado

VI Karirí family

- 1 †Kipeá/Kirirí
- 2 †Dzubukuá
- 3 †Sabuyá/Sapoyá
- 4 †Kamurú/Pedra Branca

VII Yatê family

- 1 Yatê/Fulniô/Carnijó

VIII Karajá family

- 1 Karajá

IX Ofayé family

- 1 Ofayé(-Xavánte)/Opayé

X Boróro family

- 1 Eastern Boróro
- 2 †Western Boróro
- 3 Umutína
- 4 †Otúke

XI Guató family

- 1 Guató

XII Rikbaktsá family

- 1 Rikbaktsá

† = extinct

* Two capital letter abbreviations on the map refer to the Brazilian states, e.g. *MT* stands for Mato Grosso and *MS* for Mato Grosso do Sul, etc.

Xerente (men's speech; Sinval Martins de Sousa Filho p.c.)

- (34) a. *mãĩ-p ti-kãhã?*
 what QUOTATIVE-DEM
 'What's this?'
 b. *mãĩ mã-p kmãdiki?*
 what PST-Q see
 'What did he see?'
 c. *mãĩ mã-p themba?*
 what PST-Q happen
 'What happened?'

Xerente (women's speech; Sinval Martins de Sousa Filho p.c.)

- (35) *taha-p to-tahã ambi waikud?*
 IPW QUOTATIVE-DEM man know
 'What does he (that man) know?' or 'Whom does he (that man) know?'

When speaking to men, women may use *nõkwa* 'who?', but they do not use *mãĩ* 'what?', which appears to be a men's form exclusively.

No other Jê languages has been found to have an interrogative similar to Xerente *taha* 'who?, what?'. I suppose that *taha* 'who?, what?' may be a combination of the third person singular pronoun 'he, she, it'/anaphoric demonstrative modifier 'the aforementioned [N]' *ta(-hã)*, as in (32c) and (35) (cf. also Wieseemann 1986:361, 374, 378) and the nominal *ha* 'type, kind', as in (36).

Xerente (Sinval Martins de Sousa Filho p.c.)

- (36) a. *n(h)anẽ-p ha rɔmzari?*
 how-Q **kind** animal
 'What is the kind of the animal?' (the original Portuguese translation:
 'Como é a espécie (o tipo) do animal?')
 b. *n(h)anẽ-hã-p to-kãhã rɔmzari?*
 how-**kind**-Q QUOTATIVE-DEM animal
 'What kind of animal is this?'

Thus, the original meaning of *ta-ha* may have been 'that kind, the kind'. In other words, it may have originated as a kind of filler (placeholder) 'that one, you must know...', which has become conventionalized as the interrogative 'who?, what?' in women's speech.

8.10.2 Krenák

Krenák/Krenak/Botocudo/Borun is a moribund language still spoken only by a small group of people living on the left bank of Rio Doce in the state of Minas

Gerais, Brazil, between the towns of Resplendor and Conselheiro Pena.³⁵ According to Rodrigues (1999b:167), Krenák comprised several distinct groups: Nakrehé, Nakpié, Naknyanúk, Nakyapmã, Nyepnyep, Etwet, Minyãirún, Yiporók, Pojixá, Potén, Krekmún, Bakuén, and Aranã. Together with the extinct Guerén language Krenák constitutes the Krenák family of the Macro-Jê stock.

There is a disagreement between the sources on the forms and meanings of the interrogative pronominals. Thus, according to Lucy Seki (p.c.), who has worked with the remaining speakers of Krenák, Krenák distinguishes *inãŋ* ‘who?’ and *?amnĩm* ‘what?’. However, according to Ehrenreich (1896:625), who used some older sources and his own data collected with a Krenák group on Rio Doce,³⁶ Krenák has a single interrogative *hokonim* ‘who?, what?’, as in (37).

Krenák

- (37) a. *hokonim huk nĩnum a-taĩ?*
 IPW his arm broke
 ‘Who broke his hand?’ (Ehrenreich 1896:617)
- b. *hokonim akkorune?*
 IPW 2SG.want
 ‘What do you want?’ (Ehrenreich 1896:626)

Moreover, *hokonim* can be used in the locative meaning ‘where from?’, as in (38).

Krenák (Ehrenreich 1896:616)

- (38) *hokonim antšuk nerin?* [*nĩgaram nerin uĩam*]
 IPW 2PL have.come
 ‘Where have you come from? [We have come this way.]’

The other locative interrogative is *hökrē* ‘where?’, as in (39).

Krenák (Ehrenreich 1896:615)

- (39) a. *hökrē Marão Krak inuk brauĩ?*
 where PROP mountains its way
 ‘Where is the way to the Marão mountains?’

Provided Ehrenreich’s (1896) description is correct, the lack of differentiation between ‘who?’ and ‘what?’ with the interrogative *hokonim* is most likely to be due to its origin in some locative interrogative (such as ‘where from?’ in (38)),

³⁵ Cf. <http://www.socioambiental.org/pib/epi/krenak/krenak.shtm>.

³⁶ “Ich selbst hatte auf meiner Reise am Rio Doce 1884/85 Gelegenheit, die bisher vorhandenen Wörterverzeichnisse zu kontrolliren und zu vervollständigen, konnte auch einiges grammatische Material beibringen” (Ehrenreich 1896:607).

which could also be used selectively ‘which one?’ and was subsequently extended to non-selective contexts. The interrogative *hokonim* seems to have the same root as the interrogative *hōkrē* ‘where?’, while the final part of *hokonim* resembles the final part of the demonstrative *tokonim* (Ehrenreich 1896:625).

8.10.3 Maxakalí

Maxakalí is spoken by some 800 people living in the municipality of Bertópolis of the state of Minas Gerais, Brazil (Popovich & Popovich 2005:iv). Together with several extinct languages it constitutes the Maxakalí family of the Macro-Jê stock. The data on Maxakalí interrogatives at my disposition come from Popovich (1984, 1986), Popovich & Popovich (2005) and Araújo (2000 via GICLI Swadesh list for Maxakalí). None of the sources provides a description of the interrogative pronominals of Maxakalí, so that all the data comes from examples and citation forms. Popovich & Popovich’s (2005) dictionary of Maxakalí surprisingly lacks the entry for ‘who?’, there is only an entry for the interrogative pronominal *pute(p)* ‘what?’, as in (40). The GICLI Swadesh list for Maxakalí also gives *pute* as ‘what?’.

Maxakalí

- (40) a. *pute(p) mūn mīy?*
 what 3EMPH do
 ‘What did he do?’ (Popovich & Popovich 2005:37)
- b. *’ūga putep tu kux pe mīp xax tu?*
 oh what at/to finish or wood bark at/to
 ‘[The moon having learned, went to strip off some bark. He got some *tokoxuk* tree bark. Instead of some heavy bark, he got some thin bark that arrows can go right through.] What good is that for a head protection?’ (?lit.: ‘Oh, to what will it lead, with such (thin) wood bark?’) (Popovich & Popovich 2005:37)

Furthermore, there is an “emphatic” version of ‘what?’ *pute ūm*, as in (41). The *ūm* in *pute ūm*, is identical to the modifier *ūm* ‘some, any, which/what [N]?’ as in (42).

Maxakalí

- (41) *pute ūm nūhū?*
 what (?any) DEM
 ‘But what is this?’ (Popovich & Popovich 2005:37)
- (42) a. *tík ūm te mām xuk*
 man some A fish catch
 ‘Some Maxakalí man caught the fish’ (Popovich & Popovich 2005:44)

b. *hāmpakut ũm nūhū?*

sickness which DEM

‘Which sickness is this?’ (Popovich & Popovich 2005:6)

The interrogative ‘who?’ is given as *?ũ?ũm* in the GICLI Swadesh list for Maxakalí. Popovich (1986:356) gives examples (43a) and (43b), where he glosses *ũm* and *ũ’ũm* as ‘who?’ respectively.

Maxakalí (Popovich 1986:356)

(43) a. *te ũm nūn? ũgmūn*
(?) who come 1SG.EMPH

‘Who came? I did.’

b. *ũ’ũm te-’ xupxet? hamūn*
who A-3 steal 3PL.EMPH

‘Who stole it? They did.’

The element *te* in (43a) is glossed by Popovich as the “transitive subject marker”. However, *te* cannot be such a marker here, given that according to Popovich’s (1986) own description, this marker is a postposition to the transitive subject (cf. also Popovich & Popovich 2005:iv-v), whereas in (43a) it precedes the subject of an intransitive verb. Rather, *te ũm* here seems to be the emphatic form of the interrogative *tep* ‘where?, why?’ (Popovich 1986:353), with the loss of the final *p*, similarly to *pute ũm* from *putep*.

In addition, one also finds examples like (44a) and (44b), where *putep* ‘what?’ and *tep* ‘where?, why?’ are translated as the human interrogatives ‘whose?’ and ‘who?’ respectively.

Maxakalí

(44) a. *putep koxuk nūhū?*
who image DEM

‘Whose image is this?’ (Popovich & Popovich 2005:19)

b. *tu “tep te xa ’-xaha?” kaxĩy ’axa’*

and.SAME.SUBJECT who A 2SG 3-hoe END.OF.QUOTE HEARSAY

‘[His brother, the moon came and saw him]. “Who chopped down the growth for you?” asked the moon. [The sun replied, “I teach you but you will know nothing because you do not learn.”]’ (Popovich 1984, text D.1, sentence 15)

However, the two last examples are not fully “waterproof”. Thus, the first example (44a) is presented out of context, so that it is impossible to control the translation provided in the source. The human translation of the second example (44b) sounds somewhat strange in this particular context. Given the answer ‘I teach you...’, the question would rather be expected to be ‘How did you chop

down the growth?’. Summing up, more data is needed.

8.10.4 Fulniô/Yatê

Fulniô/Yatê/Iatê/Carnijó, constituting in itself a family within the Macro-Jê stock, is spoken by some 3000 people living in the municipality of Águas Belas of the state of Pernambuco, Brazil.³⁷ The data on Fulniô interrogatives at my disposition come from the description in Lapenda (1986). Lapenda (1986:129-133) gives distinct interrogatives ‘who?’ and ‘what?’, the basic forms being *naxi* ‘who?’ (“quem?, que pessoa?”) vs. *tô* ‘what?, which one?’ (“que?, qual?, que coisa?”), *to-n-kya/tó-n-kya-* ‘which (one)?, what?’ (“qual?, que?”), *to-t(e)* ‘what?, what/which [N]?’³⁸ respectively. At the same time, at least predicatively, the interrogative base *tô/tó/to* appears to be used in questions about humans as well. In this respect, compare (45a) with (45b) and (45c).

Fulniô (Lapenda 1986:130)

- (45) a. *to-n ana:’doasê?*
 IPW-EMPH the.one.that.you.saw
 ‘What did you see?’ (lit.: ‘What is the one that you saw?’)
- b. *to-n owá seti tet’hô?*
 IPW-EMPH DEM house the.one.who.builds
 ‘Who builds this house?’ (lit.: ‘Who is the one who builds this house?’)
- c. *to-k-sa-he txua makhai?*
 IPW-for-POSS-EMPH DEM bow
 ‘Whose is that bow?’

Lapenda (1986:130) translates (45b) and (45c) using Portuguese *quem* ‘who?’. However, given that the examples are presented out of context, the selective interpretation of the interrogatives as ‘which one? (person)’ cannot be excluded.

8.11 Guaicuruan languages

The Guaicuruan/Waikurúan family includes four living members, Mocoví, Pilagá, Toba, and Kadiwéu, and one extinct language, Abipon. The South Guaicuruan languages Mocoví, Pilagá and Toba are (and Abipon was) spoken in the so-called Gran Chaco region, which covers the lowlands of southeastern Bolivia, western Paraguay and the adjacent border areas in the north of

³⁷ Cf. <http://www.socioambiental.org/pib/epi/fulnio/fulnio.shtm>.

³⁸ The final *-t(e)* in *to-t(e)* ‘what?, what/which [N]?’ appears to come from the postposition *-t(e)* ‘with’ (Lapenda 1986:134-135), even though Lapenda himself (1986:130) explicitly rejects this analysis without any further explanation.

Argentina. The only North Guaicuruan language Kadiwéu, is spoken further to the north in the state of Mato Grosso do Sul, Brazil, near the border with Paraguay and Bolivia. Guaicuruan languages are sometimes believed to constitute one larger family together with the Matacoan languages (cf. Section III.8.12).³⁹ However, the presumed relation between the two families does not appear to be commonly accepted.

All the four extant Guaicuruan languages appear to allow a lack of differentiation between ‘who?’ and ‘what?’, cf. Table 4. As to the extinct Abipon, the only source on this language at my disposition is Najlis (1966), who unfortunately does not provide enough information on the interrogative pronominals.

Table 4. Guaicuruan languages reported to have a ‘who?, what?’ interrogative

Forms & Sources	
Mocoví	<i>ñige</i> (+ DEM (“deictic classifier”), cf. below) ‘who?, what?, which one?, how?’ (Grondona 1998:162-163)
Toba	<i>nege</i> (+ DEM (“deictic classifier”), cf. below) ‘who?, what?, which one?’ (Klein 2001:23; Messineo 2003:120) <i>igá/ygadé</i> ‘who?, what?, which one?’ (Bárcena 1893:88)
Pilagá	<i>nae</i> ‘where?’ + DEM (“deictic classifier”, cf. below) = ‘who?, what?, which one?’ <i>čaqá</i> + DEM (“deictic classifier”, cf. below) = ‘who?, what?, which one?’ (Vidal 2001:346-348)
Kadiwéu	<i>ame</i> + <i>i-</i> M ⁴⁰ + DEM (“deictic classifier”, cf. below) = ‘who?, what?, which one?’ (Griffiths 1976:37-38; Sandalo 1997:73-74, 78-79)

In Mocoví and Toba, the interrogatives provided in Table 4 are normally followed by the so-called “deictic classifier”, most typically the “absent” one (also called distal, indefinite, unknown, etc.), such as Mocoví *ka* in (46) (Grondona 1998:162-163) or Toba *ka* in (47a) (Messineo 2003:120).

³⁹ This is also the point of view adopted by the Ethnologue.

⁴⁰ Structurally, the masculine marker is a prefix to the demonstrative (cf. Sandalo 1997:58-59).

Mocoví (South Guaicuruan; Argentina; Grondona 1998:162)

- (46) a. *ñige? ka i-apongi a-ñi*
 IPW DEM.ABSENT 3ACTIVE-cover F-DEM.NON<EXTENDED>
xwan l-e?ya?
 PROP 3POSS-hole
 ‘Who covered Juan’s hole?’
- b. *ñige? ka r-iyak-ir?*
 IPW DEM.ABSENT 2ACTIVE-bring-HON
 ‘What did you bring?’

Toba (South Guaicuruan; Argentina, Paraguay, Bolivia; Messineo 2003:120)

- (47) a. *nege ka?*
 IPW DEM.ABSENT
 ‘What’s the matter?’
- b. *nege na na siyagawa?*
 IPW DEM.PROX DEM.PROX man
 ‘Who is this man?’

In Pilagá and Kadiwéu, the deictic classifier appears to be obligatory for the interrogatives meaning ‘who?, what?, which (one)?’ (48-49).

Pilagá (South Guaicuruan; Argentina; Vidal 2001:347)

- (48) a. *nae-ga’ n-set da’ qad-potane?*
 where-DEM.ABSENT 3.“SET B”-dare DEM.VERTICAL 1POSS-guardian
 ‘Who dares to be our guardian?’
- b. *nae-ga’ awa-pyag-ek na’ lapat?*
 where-DEM.ABSENT 2.“SET A”-cut-OUTWARDS DEM.PROX meat
 ‘What do you cut meat with?’
- c. *nae aw-čiyof-o-ge?*
 where 2.“SET A”-descend-THITHER
 ‘Where do you come from?’

Kadiwéu (North Guaicuruan; Brazil; Sandalo 1997:73)

- (49) *ami:n:a ika ane enagi?*
 |ame-i-n:a i-ka ane y-ana-g|
 IPW-M-DEM.PROX M-DEM.ABSENT REL 3SG.S-come-TELIC
 ‘Who/what is this who/what is coming?’

The lack of differentiation between ‘who?’ and ‘what?’ in Guaicuruan languages appears to be due to the fact that the respective interrogative pronominals are originally selective interrogatives ‘which one? (person or thing)’ based on a locative interrogative ‘where?’, typically in combination with a deictic classifier.

The latter is either left unmarked for gender or, as in Kadiwéu (49), is obligatory marked as masculine, the masculine thus being the functionally unmarked gender as compared to the feminine.⁴¹

The locative/selective origin of the ‘who?, what?’ interrogatives is particularly clear in Pilagá, where (i) one of the two ‘who?, what?, which one?’ interrogatives is composed of *nae* ‘where?’ and a deictic classifier, as illustrated in (48) above, (ii) the other ‘who?, what?, which one?’ interrogative *čaqa* clearly goes back to *naečaqa* ‘which [N]?’ (cf. Vidal 2001:346-347), as in (50).

Pilagá (Vidal 2001:347)

- (50) *naečaqa-ga’ ganaʃat ga’ aw-pyag-aʃan-na?*
 which-DEM.ABSENT knife DEM.ABSENT 2. “SET A”-cut-VALENCY-INS
 ‘Which knife did you cut it (e.g., the meat) with?’

Note that the interrogative *naečaqa* itself is obviously based on *nae* ‘where?’ as well. The last part *-ča(qa)* can be compared to the distal demonstrative root *-ča*, as in *ha-da-ča* F-DEM.VERTICAL-DISTAL ‘that feminine standing’ (cf. Vidal 2001:122-124).

The Mocoví and Toba ‘who?, what?’ interrogatives, *ñige* and *nege* respectively, strongly resemble Pilagá *nae* ‘where?’. In this respect, note that whereas Messineo (2003:120) gives Toba ‘where?’ as *waʔage*, an older source,⁴² Bárcena (1893:88), gives Toba ‘where?’ as *menagé*. The latter form may also provide a link to the Kadiwéu root *ame* as found in the Kadiwéu ‘who?, what?’ interrogative [*ami*:- (< *ame* + *i*-M) + DEM] and the Kadiwéu interrogative *igame*, used to question “adjuncts”, viz. ‘why?, how?, where?’ (Sandaló 1997:73-74). The Kadiwéu *igame* also resembles Bárcena’s (1893:88) Toba interrogative *igá* ‘who?, what?, which one?’. The final *-me* may be cognate to the Pilagá medial demonstrative root *-m’e* (Vidal 2001:122-124), thus being functionally similar to *-ča* in Pilagá *nae-čaqa* ‘which [N]?’.

8.12 Matacoan languages

The Matacoan/Mataguayan languages are spoken in the same Gran Chaco region, spanning over northern Argentina, western Paraguay and southeastern Bolivia, as

⁴¹ Guaicuruan languages have two genders, masculine and feminine, which both contain animates and inanimates. Gender is rarely overtly marked on nouns, but may be marked on deictics and adjectives. However, often the masculine agreement pattern does not have any special marking and the use of the feminine agreement pattern may be optional, as for instance in Pilagá (cf. Vidal 2001:122).

⁴² It may also be possible that Bárcena (1893) and Messineo (2003) discuss different dialects of Toba.

the Guaicuruan languages discussed in Section III.8.11. As already mentioned, the two families are sometimes considered to form a higher genetic unit, the so-called Mataco-Guaicuru family (cf. Section III.8.11). Depending on what counts as a language, the Matacoan family includes four to ten languages: (i) Nivacle/Chulupí-Ashlushlay/Ajlujlay (two varieties), (ii) Chorote/Chorotí (with two varieties, treated as languages in the Ethnologue, viz. Iyo'wujwa Chorote and Iyojwa'ja Chorote), (iii) Maká/Macá/Maca (two varieties), (iv) Mataco/Wichí (Lhamtés)/Weenhayek (with four varieties, three of which treated as separate languages in the Ethnologue, viz. Wichí Lhamtés Vejoz, Wichí Lhamtés Nocten, and Wichí Lhamtés Güisnay).

I have been able to consult data on Maká, Wichí Lhamtés Nocten, and Wichí Lhamtés Vejoz. Only the first two languages appear to allow a lack of differentiation between 'who?' and 'what?'. Thus, according to Gerzenstein (1994:165-170, 178), Maká uses the interrogative *lek* 'who?, what?', frequently followed by the masculine form of the demonstrative *p-* "absent and unseen before by the speaker" [i.e., absent and unknown to the speaker] ("ausente pero no visto antes por el hablante", vs. *k-* "ausente pero visto antes por el hablante"),⁴³ as in (51).

Maká (Matacoan; Paraguay; Gerzenstein 1994:178)

- (51) a. *lek p-a'* *ne-fen?*
 IPW DEM.ABSENT&UNKNOWN-M 3>2-help
 'Who helped you?'
 b. *lek p-a'* *Ø-tux n-a'* *k'utsax?*
 IPW DEM.ABSENT&UNKNOWN-M 3A-eat DEM.MEDIAL-M old.man[M]
 'What did the old man eat?'

Lindström (1995:313) citing Kenneth Claesson (p.c.) reports that Wichí Lhamtés Nocten (the Bolivian variety of Mataco/Wichí Lhamtés) allows a lack of differentiation between 'who?' and 'what?' with the interrogative *háats'i*, as in (52).

Wichí Lhamtés Nocten (Matacoan; Bolivia)

- (52) *háats'i tà ya-hoo-Ø-yeh?*
 IPW COMP 3S-go-3OBJ-to
 'Whom/what did he go to?', 'Who/what went to him?', 'Who/what went to it?' (Lindström 1995:313 citing Kenneth Claesson, p.c.)

It is also possible to use *háats'i* as a modifier 'which [N]?', 'what [N]?' to the nouns meaning 'person' and 'thing', thus explicitly distinguishing between

⁴³ In this respect, recall the same use of the masculine absent demonstrative with the 'who?, what?' interrogatives in Guaicuruan languages (Section III.8.11).

‘who?’ (lit.: ‘what person?’) and ‘what?’ (lit.: ‘what thing?’). The “preferred interpretation of *háats’i* is often human, but it depends a lot on context” (Lindström 1995:313). Furthermore, the locative and temporal interrogative meanings are more frequently expressed with the specialized interrogative *kyi*, rather than by the combination of *háats’i* with a locative (or temporal) marker, such as *yeh*.

In the closely related Wichí Lhamtés Vejoz, as described by Viñas Urquiza (1974:106-107), the cognate interrogative *áci* when used pronominally (“función pronominal”) may mean only ‘who?’ (53a). It can also be used attributively (“función determinativa en contexto de sustantivo”) as ‘what [N]?, which [N]?’. The non-human meaning ‘what?’ is expressed by the phrase *áci mak* ‘what thing?’ (53b). The locative interrogative ‘where?’ has the form *čĩ*.

Wichí Lhamtés Vejoz (Matacoan; Argentina)

- (53) a. *áci ta hip’oámhiče?*
 IPW COMP 3>2.shut.in
 ‘Who shut you in?’ (Viñas Urquiza 1974:107)
- b. *áci mák ta lew’atla?*
 IPW thing COMP 2.need
 ‘What do you need?’

9 Lack of differentiation: concluding remarks

9.1 Main results

In the preceding sections, I have discussed various languages that (may) allow for a lack of differentiation between ‘who?’ and ‘what?’, as well as the possible origins (if any) of this lack of differentiation in each particular case. The main results can be summarized as in (1).

- (1) a. Differentiation between ‘who?’ and ‘what?’ is not universal and although on a world-wide scale, lack of differentiation between ‘who?’ and ‘what?’ is not common, it is far from being as marginal as is often assumed.
- b. Languages allowing for a lack of differentiation between ‘who?’ and ‘what?’ can be found all around the world. However, the distribution is rather uneven, which points to areal influences and above all, genetic predispositions.
- c. Languages may both lose differentiation between ‘who?’ and ‘what?’ and acquire it. In many cases, it has proved possible to trace the origins of the lack of differentiation between ‘who?’ and ‘what?’ in a given language. Several paths appear to recur in unrelated and geographically distant languages.

In what follows, I will discuss these results in more detail, viz. (1a) in Section III.9.2, (1b) in Section III.9.3, and (1c) in Section III.9.4.

Before proceeding, it may be useful to recall what kinds of lack of differentiation between ‘who?’ and ‘what?’ can be distinguished. As discussed in Section I.4, besides an “unrestricted” lack of differentiation (Section I.4.1), two further possibilities are as follows: (i) the lack of differentiation between ‘who?’ and ‘what?’ is restricted to predicatively used non-selective interrogative pronominals (Section I.4.2.1) and/or (ii) the lack of differentiation between ‘who?’ and ‘what?’ is restricted to a certain gender and/or number (Section I.4.2.2). What is more, in the former case, there appears to be an implicational relation, cf. (2).

- (2) predicative < non-predicative

That is, if ‘who?’ and ‘what?’ must be distinguished in non-predicatively used non-selective interrogative pronominals, then also in predicative use, whereas the opposite is not necessarily the case. Judging from the data available, the implicational relation in (2) is (nearly?) universal. The only possible exception I know of might be constituted by Old Telugu (South-Central Dravidian; India; cf. note ^{vii}) to the table in Appendix E), but I do not have the relevant data to make

clearer statements in this respect.

Finally, it is worth reminding that what matters most for me when discussing possible cases of a lack of differentiation between ‘who?’ and ‘what?’ is whether a given language *allows* for a lack of differentiation between ‘who?’ and ‘what?’. Whether the language requires this or not is of secondary importance.

9.2 Lack of differentiation between ‘who?’ and ‘what?’: how widespread it is and why

As has been demonstrated in the present study, differentiation between ‘who?’ and ‘what?’ is not universal. In my global sample of some 1850 languages, approximately 7-9% of the languages have been found that (may) allow a lack of differentiation between ‘who?’ and ‘what?’ (cf. Section III.9.3 for a discussion of the distribution of these languages). My use of only approximate numbers here is due to several reasons. To begin with, the number 1850 itself is approximate because it consists of 1048 idioms of my own database and approximately 470 extra Bantu and 330 extra Austronesian idioms from the databases of Bastin *et al.* (1999) and Blust *et al.* (2006) respectively. I have not fully incorporated the latter two databases in my own database. Second, in some languages (131 languages) the case for a lack of differentiation appears to be more solid than in others (37 additional languages). That is, the higher percentage refers to both kinds of languages together. I have rounded the percentages to whole numbers. Third, there may be some languages in my sample allowing for a (at least “restricted”) lack of differentiation between ‘who?’ and ‘what?’ (cf. Section III.9.1) which I have overlooked because this information is not mentioned explicitly in the source I consulted. Finally, it is not always easy to distinguish consistently between a language and a dialect, i.e. what some would count as a separate language, others may prefer to treat only as a dialect. Differences may be quite drastic here (cf., e.g., Section III.6.2 on Nyulnyulan languages). However, counting languages may cause problems even when everybody agrees in treating given idioms as dialects of one language rather than distinct, albeit closely related languages. Thus, whereas some dialects may require a differentiation between ‘who?’ and ‘what?’, others may lack it. In such cases, I usually counted only dialects that allow for a lack of differentiation between ‘who?’ and ‘what?’.

We may now try to answer two related questions. First, why do (most) languages oblige their speakers to differentiate between ‘who?’ and ‘what?’? Or from a different angle: How is it possible that (some) languages do not oblige their speakers to differentiate between ‘who?’ and ‘what?’, given the fundamental nature of the distinction between persons and things as the basic means of human categorization of entities? The answer to the second question seems to be quite straightforward. If I ask a constituent question, which is based

on the presupposition *It is known that (possibly) HAPPEN/EXIST (...x...)* (cf. Section I.2.2), it means that I do not know what variable *x* is. Although I may have some ideas as regards the kind of entity that *x* belongs to, in principle I should not be expected to do so. Thus, languages that allow for a lack of differentiation between ‘who?’ and ‘what?’ do not require this from their speakers.

The first question is more interesting and less obvious than it may seem at first sight. Indeed, if there is no need for the speaker to know what kind of entity variable *x* belongs to, why should most languages bother their speakers with this issue by obliging them to distinguish between ‘who?’ and ‘what?’? To begin with, all languages differentiating between ‘who?’ and ‘what?’ appear to dispense their speakers from the need to choose between ‘who?’ and ‘what?’ when the speaker really does not have the slightest idea of whether the entity in question is a person or a thing (or for one or another reason, does not wish to make any conjecture on this point). As discussed in Section I.4.2.3.3.1, in such “no conjecture”-context, the languages of the world appear to be just like English in preferring the interrogative ‘what?’, which is thus the default option. Speakers are obliged to choose between ‘who?’ and ‘what?’ only in contexts, which are by far more numerous than “no conjecture”-contexts, when they are able to make at least some assumptions on the kind of entity that variable *x* belongs to. Given the fundamental nature of the distinction between PERSON and THING as a means of human categorization of entities (cf. Section I.2.4), it seems only natural that this distinction would be used by the speakers as the opposition in terms of which they would construe their first assumptions on the kind of entity that variable *x* belongs to. However, even though all this may account for the fact that the opposition between ‘who?’ and ‘what?’ is not unexpected and that in principle, it can be rather common, this does not explain why it is so widespread.

I suppose we may find an answer to this question if we compare non-selective interrogative pronominals with such elements as selective interrogative pronominals, anaphors, demonstratives and indefinites. For instance, Diessel (2003:641) notes that “the distinction between the features person and thing is much more common in interrogatives than in demonstratives”. Not surprisingly, selective interrogative pronominals, which typically expect a demonstrative as an answer, appear to resemble demonstratives in this respect (cf. Sections I.2.3, I.2.6). Similarly, I have the strong impression that the distinction between the features person and thing in anaphoric pronominals is also rare, although perhaps it may be somewhat more frequent as compared to demonstratives. On the contrary, with indefinite pronominals (at least some kinds of), the distinction between the features person and thing seems to be as common as it is with non-selective interrogative pronominals. This similarity between indefinites and interrogatives is not particularly surprising, since many other formal and

semantic parallels between the two categories are known to exist (cf., e.g., Mushin 1995:1-6, Haspelmath 1997:174-176 and Bhat 2000, 2004:226-249, also see there for further references).

Diessel (2003:647) is probably correct when he says that both demonstratives and interrogatives, and we may also add here anaphors and (with some reservations) indefinite pronominals, “can be seen as signals that instruct the hearer to search for a specific referent”, they “initiate a search for information that is guided by their semantic features”. However, demonstratives, anaphors and indefinites/interrogatives appear to differ substantially in the kind of search they initiate. Thus, demonstratives, which are basically exophors, initiate a search for a specific referent in the space surrounding the interlocutors. In other words, the only thing that an exophoric element must do is to pick a referent out of the surrounding space by referring to its location in relation to some deictic centre. Anaphoric pronominals initiate a search for a specific mentioning of a referent in discourse (usually the preceding discourse) and thus indirectly to the referent itself. Clearly, the accessibility of the referent here is not as immediate as it is in the case of exophoric reference. Finally, the search that is initiated by interrogatives (and to a large extent indefinites) is directed neither to a (preceding or future) mentioning of a referent nor, strictly speaking, to the referent itself. Rather, the search is for a word (or a combination of words), viz. a linguistic sign (in this particular case, a nominal), that can be used to refer to the referent. This is particularly clear in the case of questions like *Who is this?* or *What is this?*, when the referent is immediately present and accessible to both the speaker and the interlocutor. Summing up, while demonstratives have to do with locating a referent in the surrounding space and anaphors have to do with locating a word form in the discourse used to refer to the referent, interrogatives (and some indefinites) have to do with finding a word (or words) in the lexicon that can be used to refer to the referent. In other words, we observe here a gradual strengthening of the association with the lexicon: absent in the case of demonstrative nominals, only indirectly present in the case of anaphoric pronominals and directly present in the case of non-selective interrogative pronominals (and some indefinites). At the same time, the lexicon, in its nominal domain, reflects human categorization of entities and as discussed in Section I.2.4, the distinction between persons and things is fundamental to the latter. The two facts considered together, viz. (i) the direct association of non-selective interrogative pronominals (and some indefinites) with the lexicon and (ii) the fundamental nature of the distinction between persons and things in the (nominal) lexicon, appear to account rather straightforwardly for the fact that the opposition between persons and things in the non-selective interrogative pronominals, i.e. the opposition between ‘who?’ and ‘what?’, is so widespread in the languages of the world. The question why it does not need to be universal has

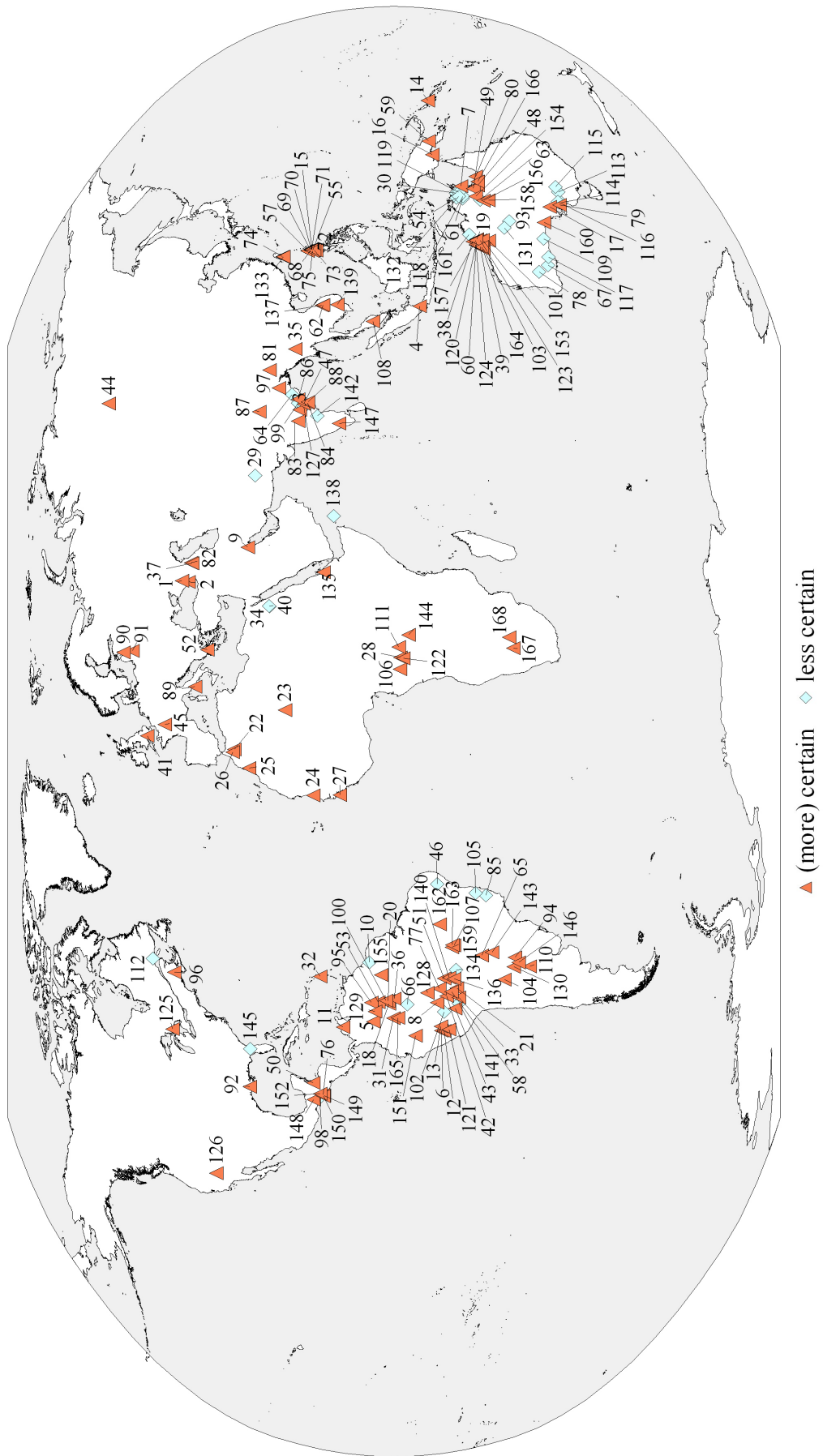
been considered earlier in this section.

9.3 Lack of differentiation between ‘who?’ and ‘what?’: areal and genetic distribution

As already mentioned in Section III.9.2, out of my global sample of some 1850 languages, approximately 7-9% of the languages have been found that (may) allow a lack of differentiation between ‘who?’ and ‘what?’. Taking into account the reservations made in Section III.9.2, we may speak of 131 rather solid and at least 37 further possible cases of a lack of differentiation between ‘who?’ and ‘what?’. The areal distribution of these languages is illustrated on Map 1. As can be observed, although languages allowing a lack of differentiation between ‘who?’ and ‘what?’ can be found all around the world, their distribution is rather uneven. However, it is somewhat difficult to decide how to pin down the unevenness of this distribution in strict numerical terms. There are two main problems here. First, what should we count: languages allowing for a lack of differentiation between ‘who?’ and ‘what?’ or genera with such languages,¹ for instance? Second, should the results of different areas be compared to each other in absolute terms or should we count relative percentages by area and compare these to each other.

¹ A genus is “a genetic group roughly comparable to the subfamilies of Indo-European” (Dryer 2003:110), “of a time depth of 3,500 to 4,000 years” (Dryer 2000:335). Note, though, that in many cases a genus corresponds just to the deepest universally accepted level of grouping of related languages.

Map 1. The distribution of the languages that (may) allow a lack of differentiation between 'who?' and 'what?'



Language	Family (Ethnologue)	Language	Family (Ethnologue)	Language	Family (Ethnologue)
1 Abaza (Tapanta)	North Caucasian	22 Berber ("continuous" Zenati, Aït Seghrouchen of Oum Jeniba)	Afro-Asiatic	41 English (Old)	Indo-European
2 Abkhaz	North Caucasian	23 Berber (Ahaggar Tuareg)	Afro-Asiatic	42 Ese Eija/Essejja (Bolivia)	Tacanan
3 Abujmaria	Dravidian	24 Berber (Mauritanian)/ Zenaga	Afro-Asiatic	43 Ese Eija/Essejja (Peru, Tacanan)	Tacanan
4 Abung/Lampung Nyo (Melinting)	Austronesian	25 Berber (Tachelhit)	Afro-Asiatic	Baawaja Kuñaje of Tambopata river)	
5 Achagua	Arawakan	26 Berber (Tamazight, Ayt Ndhir)	Afro-Asiatic	44 Evenki (Western, Altaic)	Altaic
6 Ajyíninka (Apurucayali)	Arawakan	27 Bidyogo/Bijogo (Kagbagaa)	Niger-Congo	45 French	Indo-European
7 Anindilyakwa	Australian	28 Bolia	Niger-Congo	46 Fulniô/Yatê*	Macro-Ge
8 Apurinã	Arawakan	29 Brahui*	Dravidian	47 Gadaba	Dravidian
9 Arabic (Mesopotamian)	Afro-Asiatic	30 Burarra*	Australian	48 Garawa/Garrwa (Eastern)	Australian
10 Arawak/Lokono*	Arawakan	31 Cabiari	Arawakan	49 Garawa/Garrwa (Western)	Australian
11 Arhuaco/Ika	Chibchan	32 Carib (Island, Dominica)	Arawakan	50 Garifuna/Central Ameri-	Arawakan
12 Ashéninka (Pichis)	Arawakan	33 Cavineña	Tacanan	can Carib/Black Carib	
13 Ashéninka (Ucayali-Yurúa)	Arawakan	34 Coptic*	Afro-Asiatic	(Hopkins & Livingston)	
14 Babatana (Sisingga)	Austronesian	35 Danau	Austro-Asiatic	51 Gavião do Jiparaná/ Tupi	Tupi
15 Balangao (Natonin)	Austronesian	36 Dâw	Maku	Gavião do Rondônia/ Digat/Ikôro	
16 Bamu (Pirupiru)	Trans-New Guinea	37 Dido/Tsez (Kidero)	North Caucasian	52 Greek (Ancient)	Indo-European
17 Banggarla/Parnkalla	Australian	38 Djawi	Australian	53 Guarekena/Warekena	Arawakan
18 Baniwa (of Içana)	Arawakan	39 Dyugun	Australian	54 Guragone/Gurr-goni*	Australian
19 Bardi	Australian	40 Egyptian (pre-Coptic)*	Afro-Asiatic	55 Ifugao (Bayninan)	Austronesian
20 Baré	Arawakan			56 Ifugao (Lagawe)	Austronesian
21 Baure	Arawakan			57 Isnag (Dibagat)	Austronesian

Language	Family (Ethnologue)	Language	Family (Ethnologue)	Language	Family (Ethnologue)
58 Itene/Moré	Chapacura-Wanham	82 Khvarshi (Inkhokari)	North Caucasian	104 Mataco (Wichi Lhamtés Nocten)	Mataco-Guaicuru
59 Iwal	Austronesian	83 Kolami (Kolami & Naiki)	Dravidian	105 Maxakalí*	Macro-Ge
60 Jabirjabirrr/Dyaberdyaber	Australian	84 Konda	Dravidian	106 Mbosi/Mboshi	Niger-Congo
61 Jawoyn/Djauan*	Australian	85 Krenak/Botocudo/Borun*	Macro-Ge	107 Mehinaku	Arawakan
62 Jeng/Cheng	Austro-Asiatic	86 Kui	Dravidian	108 Mintil	Austro-Asiatic
63 Jingulu/Djingili*	Australian	87 Kusunda	Isolate	109 Mirning/Mirninj*	Australian
64 Juang*	Austro-Asiatic	88 Kuvi	Dravidian	110 Mocoví	Mataco-Guaicuru
65 Kadiwéu	Mataco-Guaicuru	89 Latin	Indo-European	111 Mongo-Nkundo	Niger-Congo
66 Kaixana*	Arawakan	90 Latvian	Indo-European	112 Montagnais*	Algic
67 Kalarko/Kalaaku/Malba*	Australian	91 Lithuanian	Indo-European	113 Narrinyeri	Australian
68 Kalinga (Guinaang)	Austronesian	92 Louisiana Creole	Creole	(Keramin/Kureinji)*	
69 Kalinga (Limos) (1)	Austronesian	93 Lurija/Lurija*	Australian	114 Narrinyeri (Ngayawang)*	Australian
70 Kalinga (Lower Tanudan)	Austronesian	94 Macá	Mataco-Guaicuru	115 Narrinyeri (Yitha-Yitha/ Dardi-Dardi)*	Australian
71 Kalinga (Upper Tanudan)	Austronesian	95 Maipure	Arawakan	116 Narrinyeri (Yuyu/Ngariket)*	Australian
72 Kallahan (Kayapa Proper)	Austronesian	96 Malecite-Passamaquoddy	Algic	117 Ngadjunmaya*	Australian
73 Kallahan (Keley-i)	Austronesian	97 Malto (Malpaharia)	Dravidian	118 Ngalakan*	Australian
74 Kananabu (1)	Austronesian	98 Mam (Northern, San Pedro Necta)*	Mayan	119 Ngandi*	Australian
75 Kankanay (Northern)	Austronesian	99 Manda*	Dravidian	120 Nimanbur(ru)	Australian
76 Kaqchikel (most dialects)	Mayan	100 Mandahuaca/Mandawaka	Arawakan	121 Nomatsiguenga	Arawakan
77 Karitiána	Tupi	101 Mangarla/Mangala	Australian	122 Ntomba-Inongo	Niger-Congo
78 Karlamay*	Australian	102 Manitenére/Mantinera*	Arawakan	123 Nyikina/Nyigina	Australian
79 Kurna	Australian	103 Martu Wangka (Yulparija)	Australian		
80 Kayardild	Australian				
81 Khasi	Austro-Asiatic				

Language	Family (Ethnologue)	Language	Family (Ethnologue)	Language	Family (Ethnologue)
124 Nyulnyul	Australian	139 Stieng (Budeh?)	Austro-Asiatic	152 Uspanteka (Uspantán Centro)	Mayan
125 Ottawa/Nishnaabemwin	Algonic	140 Suruí (Paite)	Tupi	153 Walmajarri	Australian
126 Panamint/Timbisha	Uto-Aztecan	141 Tacana*	Tacanan	154 Wambaya	Australian
Shoshone/Tümpisa		142 Telugu*	Dravidian	155 Wapishana/Wapixana	Arawakan
Shoshone (Eastern)		143 Terêna	Arawakan	156 Warlmanpa	Australian
127 Parji/Duruwa*	Dravidian	144 Tetela (some dialects)	Niger-Congo	157 Warrwa	Australian
128 Paumari	Arauan	145 Timucua*	Isolate	158 Warumungu	Australian
129 Piapoco	Arawakan	146 Toba	Mataco-Guaicuru	159 Waurá	Arawakan
130 Pilagá	Mataco-Guaicuru	147 Toda	Dravidian	160 Wirangu	Australian
131 Pintupi*	Australian	148 Tzotzil (an unidentified dialect of San Andrés municipality)	Mayan	161 Worora/Worrorra*	Australian
132 Rembarunga/Rembarrnga*	Australian	149 Tz'utujil/Tzutujil (Eastern: Santiago Atitlán)	Mayan	162 Xerente/Akwê-Xerente	Macro-Ge
133 Saaroa	Austronesian	150 Tz'utujil/Tzutujil (Western: San Juan la Laguna)	Mayan	163 Yawalapiti	Arawakan
134 Sabanês*	Nambiquaran	151 Urarina	Isolate	164 Yawuru	Australian
135 Saho (Irob & Minifero)	Afro-Asiatic			165 Yucuna	Arawakan
136 Sakirabiá/Mekens	Tupi			166 Yukulta/Ganggalida	Australian
Sakirabiat				167 !Xóõ	Khoisan
137 Sapuan	Austro-Asiatic			168 #Hõã	Khoisan
138 Soqotri*	Afro-Asiatic				

* less certain

The results of different counts are summarized in Table 1. The columns represent the result per area, with the final column providing the total for all the areas. The first four rows provide the results of counting languages and the last four rows represent the results of counting genera. The first row [$N(L_{\text{lack}})$ [% of total]] gives per area numbers of languages allowing for a lack of differentiation between ‘who?’ and ‘what?’, with in square brackets their percentage in the total, which is given in the last column. The second row [$T(L)$ [% of total]] gives total numbers of languages in each area (based on Gordon 2005, excluding creoles, pidgins, artificial, sign and mixed languages), with in square brackets their percentage in the total, which is given in the last column. The third row [% of $N(L_{\text{lack}})$ in $T(L)$] represents the percentage of the number of languages allowing for a lack of differentiation between ‘who?’ and ‘what?’ (i.e., the numbers in the first row) in the total number of languages in a given area (i.e., the second row). The fourth row [C_L : > or < than average (3,32%)] gives per area results of subtraction of the average for the results for the seven areas in the third row from the percentages in the third row. The percentage in the fourth row may be either positive or negative if in a given area the result in the third row is higher than the average or lower than the average respectively. The white numbers in a black circle after each result indicate the rank of a given area² in this particular count (i.e. within the same row). In every row where languages (or genera) with lack of differentiation between ‘who?’ and ‘what?’ are counted, the area with the highest rank, viz. ❶, is highlighted in dark grey and the area with the lowest rank, viz. ❷ (or sometimes ❸), is highlighted in light grey in each row. The counts in Table 1 are done for the 131 more solid cases of lack of differentiation only. All percentages, except in the row [% of $N(L_{\text{lack}})$ in $T(L)$], are rounded to the closest whole number. Note that I counted Louisiana French Creole together with the Romance languages of Eurasia and not as a separate genus within North America. Similarly, I counted the Arawakan language Garifuna with Arawakan languages of South America, even though Garifuna is spoken in North America.

² As pointed out in Section III.1, in the present study I distinguish seven areas: (i) Africa and the Middle East, (ii) Eurasia, (iii) Southeast Asia and Oceania, (iv) New Guinea, (v) Australia, (vi) North America, (vii) South America. In most respects, my areas correspond to Matthew Dryer’s macro-areas, as found in Haspelmath *et al.* (2005), which seems to be the latest version. The only major point of difference with Dryer’s macro-areas consists in the fact that I prefer to consider New Guinea and Australia separately.

Table 1. Counting the distribution of the languages allowing for a lack of differentiation between ‘who?’ and ‘what?’

	Africa & the Middle East	Eurasia	SEAsia & Oceania	New Guinea	Australia	North America	South America	total
$N(L_{lack})$ [% of total]	15 [11%]	21 [16%]	22 [17%]	1 [1%]	23 [18%]	8 [6%]	41 [31%]	131
Ranking₁	5	4	3	7	2	6	1	(7%) ⁱ⁾
$T(L)$ [% of total]	2173 [31%]	658 [9%]	1944 [28%]	852 [12%]	263 [4%]	617 [9%]	509 [7%]	7016
Ranking₂	1	4	2	3	7	5	6	(379%) ⁱ⁾
% of $N(L_{lack})$ in $T(L)$	0,69%	3,19%	1,13%	0,11%	8,75%	1,3%	8,06%	1,87%
Ranking₃	5	3	6	7	1	4	2	
C_L : > or < than average (3,32%)?	-2,63%	-0,13%	-2,19%	-3,21%	+5,43%	-2,02%	+4,74%	
<hr/>								
$N(G_{lack})$ [% of total]	7 [14%]	12 [24%]	8 [16%]	1 [2%]	6 [12%]	3 [6%]	13 [26%]	50
Ranking₄	4	2	3	7	5	6	1	
$T(G)$ [% of total]	64 [14%]	39 [9%]	43 [9%]	92 [20%]	33 [7%]	93 [20%]	92 [20%]	456
Rankings₅	3	5	4	2	6	1	2	
% of $N(G_{lack})$ in $T(G)$	11%	31%	19%	1%	18%	3%	14%	11%
Ranking₆	5	1	2	7	3	6	4	
C_G : > or < than average (14%)?	-3%	+17%	+5%	-13%	+4%	-11%	same	

ⁱ⁾ % of the number of languages in my global sample, viz. ca. 1850

$N(L_{lack})$ number of languages allowing for a lack of differentiation between ‘who?’ and ‘what?’ in a given area

$T(L)$ total number of languages in a given area (based on Gordon 2005, excluding creoles, pidgins, artificial, sign and mixed languages)

$N(G_{lack})$ number of genera allowing for a lack of differentiation between ‘who?’ and ‘what?’ (genera counted following Haspelmath *et al.* 2005)

$T(G)$ total number of genera in a given area (based on Haspelmath *et al.* 2005)

As can be observed in Table 1, depending on the way we count, the rankings of the areas as regards the frequency of lack of differentiation between ‘who?’ and ‘what?’, viz. Ranking₁, Ranking₃, Ranking₄, and Ranking₆, may vary substantially. Interestingly, irrespective of the way we count, three areas strongly tend to get the lowest ranking, viz. (in descending order) Africa & the Middle East > North America > New Guinea.³ As to the higher end of the ranking, South America and Australia appear to do better on languages and Eurasia and Southeast Asia & Oceania on genera. However, two remarks are appropriate here. First, whereas 8 of the 12 Eurasian genera belong to just two universally accepted phyla, viz. Indo-European and Dravidian, and similarly, all the 8 genera of Southeast Asia & Oceania belong to just two universally accepted phyla, viz. Austronesian and Austro-Asiatic, in Australia and South America the situation is much less clear. The 6 Australian and 13 South American genera under consideration are not distributed among two or three universally accepted superordinate phyla. Second, Eurasia differs from the other three areas in that in Eurasia, the lack of differentiation between ‘who?’ and ‘what?’ is typically restricted to certain genders, numbers and/or predicative use (cf. Section III.9.1), whereas in the other areas, an unrestricted lack of differentiation between ‘who?’ and ‘what?’ appears to be more common.

As to the choice between counting languages and counting genera, it seems preferable to count genera. The main problem with counting languages is that as pointed out in Section III.9.2 it is often difficult to distinguish consistently between a language and a dialect. Furthermore, there are several cases where relatively large groups of closely related idioms allow for a lack of differentiation between ‘who?’ and ‘what?’. The Arawakan languages of South America probably constitute the most extreme example here. Although this shows that a given linguistic group may have a particular predisposition for a lack of differentiation between ‘who?’ and ‘what?’, counting all these closely related languages might obscure the real pattern of the distribution of lack of differentiation between ‘who?’ and ‘what?’.

As to the choice between comparison of absolute and relative counts, I give preference to the latter. The reason is that a relative count such as Ranking₆ allows us to compare the frequencies of lack of differentiation between ‘who?’ and ‘what?’ between the different areas, which tells us more about how uneven lack of differentiation between ‘who?’ and ‘what?’ is distributed on a world-wide scale than a comparison of absolute counts would do. The uneven character of the distribution of lack of differentiation between ‘who?’ and ‘what?’ is made most evident in C_L and C_G. Row C_L of Table 1 shows that Australia and South America have more languages allowing for a lack of differentiation between

³ Note, though, that such an overall minimal ranking of New Guinea may be due to the fact that New Guinean languages are somewhat underrepresented in my sample.

‘who?’ and ‘what?’ than the average (3,32%), whereas other areas have less such languages than the average. Row C_G shows that Eurasia, Southeast Asia & Oceania and Australia have more genera with languages allowing for a lack of differentiation between ‘who?’ and ‘what?’ than the average (14%), whereas South America has the same number of such genera as the average and other areas have less.

9.4 Lack of differentiation between ‘who?’ and ‘what?’: the many ways to get there

In my discussion of particular languages allowing for a lack of differentiation between ‘who?’ and ‘what?’, I have tried to demonstrate that languages may both lose differentiation between ‘who?’ and ‘what?’ and acquire it. Indirectly, this fact undermines the universality of Haspelmath’s (1997:176) claim that “interrogative pronouns are among the slowest-changing elements in *any* language” (emphasis added). The universality of this claim is further undermined by the existence of clear examples of linguistic groups in which interrogative pronominals appear to be among the fastest-changing elements. In this respect, recall for instance Arawakan and Mayan languages discussed in Section III.8.2 and Section III.7.5 respectively. Consequently, Haspelmath’s claim must be relativized: interrogative pronominals are among the slowest-changing elements in *some* linguistic groups (e.g., Indo-European languages, which Haspelmath (1997:176) cites himself in support of his claim). The question how common such groups are requires further investigation.

In many cases, it has proved possible to trace the origins of the lack of differentiation between ‘who?’ and ‘what?’ in a given language. Moreover, several patterns of the development of ‘who?, what?’ interrogatives can be distinguished that appear to recur in unrelated and geographically distant languages. These patterns are summarized in (3). After each pattern, I provide (i) references to the sections discussing the areas where the languages with this pattern occur and (ii) the names of the respective linguistic groups where such languages are found.

(3) a. **‘where?’ > ‘which one?’ > ‘who?, what?’**

Section III.2.1.1 (Bantu), Section III.2.3.5 (Cushitic), Section III.4.2 (Austronesian), Section III.6.4.2 (possibly some non-Pama-Nyungan languages), Section III.7 (Uto-Aztecan, Mayan, and tentatively, Timucua), Section III.8 (Arawakan, Guaicuruan, possibly also Arauan, Chapacura-Wanham, and Macro-Jê)

b. **‘which one?’, ‘which/what [N]?’ > ‘who?, what?’**

Section III.3.1 (Indo-European), Section III.3.4 (Kusunda), Section III.4.1 (Austro-Asiatic), Section III.4.2 (Austronesian)

c. **[single IPW-root + GENDER-NUMBER marking] + peculiarities of the gender-number system > ‘who?, what?’**

Section III.2.2 (Tuu, alias Southern Khoisan), Section III.3.1 (Indo-European), Section III.3.2 (North Caucasian), Section III.3.3 (Dravidian), Section III.4.2 (Austronesian),⁴ Section III.6.4.2 (some non-Pama-Nyungan languages), Section III.7 (Algonquian), Section III.8 (Arawakan, possibly also Arauan, Chapacura-Wanham, Guaicuruan and Matacoan)

d. **constructions based on a noun meaning ‘name’ or verbs meaning ‘do, say, be’, ‘name’, ‘call’ > ‘who?, what?’**

Section III.4.2 (Austronesian), Section III.7.5 (Mayan), Section III.8 (Arawakan, possibly also Tacanan and Urarina)

e. **? filler (placeholder) meaning ‘whatsi(t)sname’ > ‘who?, what?’**

Section III.3.5 (Tungusic), Section III.8 (Macro-Jê and Tacanan)

f. **? language contact**

Section III.2.3.4 (Semitic), Section III.3.5 (Tungusic), Section III.7 (Algonquian)

Note that several patterns, viz. (3a-c), often appear to involve the predicative use as a necessary condition for the development of a ‘who?, what?’ interrogative, which probably has something to do with the hierarchy in (2) (cf. Section III.9.1). Furthermore, several patterns, viz. (3c), (3d) and to some extent (3b), imply a tendency for the development of non-selective interrogative pronominals out of phrasal and even clausal constructions not based on such nominals as ‘person’ or ‘thing’. Obviously, such a constructional build-up of non-selective interrogative pronominals makes their association with the lexicon less prominent. At the same time, as discussed in Section III.9.2, this association is largely responsible for the fact that the opposition between persons and things in the non-selective interrogative pronominals, i.e. the opposition between ‘who?’ and ‘what?’, is so widespread in the languages of the world. Finally and probably most importantly, in (almost) all patterns the elements which serve as the source for ‘who?, what?’ interrogatives are themselves (largely) indifferent to the distinction between person and thing.

⁴ In the case of Austronesian languages, what is involved is not gender-number marking, but a so-called “noun phrase marker” **si*.

Appendix A

The questionnaire respondents

In what follows, I provide a list with the names of the people who generously made time to answer (entirely or partially) my questionnaire for a given language. The list is organized by the names of the respective languages.

Language	Family (Ethnologue)	Person(s)
Alawa	Australian	Margaret C. Sharpe
Albanian	Indo-European	Elton K. Prifti
Alutor (Proper)	Chukotko-Kamchatkan	Yukari Nagayama
Ambai	Austronesian	Tamara Price
Amele	Trans-New Guinea	John Roberts
Amharic	Afro-Asiatic	Seyoum G. Mulugeta, Joachim Crass
Andamanese (Great)	Andamanese	Abhishek Avtans
Apali	Trans-New Guinea	Martha L. Wade
Arabela	Zaparoan	Rolland G. Rich
Arabic (Chadian)	Afro-Asiatic	Madeleine Somte
Arapaho	Algic	Andrew Cowell
Arawak/Lokono	Arawakan	Marie-France Patte
Armenian (Eastern)	Indo-European	Hratchik Martirosyan, Jasmine Dum-Tragut
Awara	Trans New-Guinea	Susan Quigley
Awetí	Tupi	Sebastian Drude
Awtuw	Sepik-Ramu	Harry Feldman
Aymara	Aymaran	Martha J. Hardman
Azerbaijani (North)	Altaic	Gilles Authier
Badaga	Dravidian	Christiane Pilot-Raichoor
Badyara/Badiaranke	Niger-Congo	Rebecca Cover
Bafanji	Niger-Congo	Cameron Hamm
Bambara	Niger-Congo	Valentin Vydrine
Barasana	Tucanoan	Elsa Gómez-Imbert
Baruga (Tafota)	Trans-New Guinea	Cindi Farr
Basque	Isolate	Itziar Laka Mugartza
Batak (Karo)	Austronesian	Geoff Woollams
Baure	Arawakan	Swintha Danielson
Beja/Bedawi (Hadendowa)	Afro-Asiatic	Martine Vanhove, Mohamed-Tahir Hamid Ahmed
Bengali	Indo-European	Aditi Lahiri, Tanmoy Bhattacharya

Language	Family (Ethnologue)	Person(s)
Berber (Ahaggar Tuareg)	Afro-Asiatic	Karl-G. Prasse
Berber (Kabyle, Tashelhit)	Afro-Asiatic	Lionel Galand
Berber (Zenaga)	Afro-Asiatic	Catherine Taine-Cheikh
Bicolano (Central, Legaspi)	Austronesian	Veronika Mattes
Bidyogo/Bijogo	Niger-Congo	Guillaume Segerer
Bisa	Niger-Congo	Colette Bancé
Bislama	Creole	Claudia Gerstner-Link
Bontok (Central, Guinaang)	Austronesian	Lawrence A. Reid
Boruca	Chibchan	Juan Diego Quesada
Budukh	North Caucasian	Gilles Authier
Bulgarian	Indo-European	Petar Kehayov
Buriat (Russia)	Altaic	Galina Dyrxeeva, Julija Badmaeva
Burmese	Sino-Tibetan	Justin Watkins, Nicoletta Romeo, John Okell
Burushaski	Isolate	Steve Willson
Candoshi	Isolate	John Tuggy
Canela-Kraho/Krahô	Macro-Ge	Jack Popjes
Capanahua	Panoan	Eugene Loos
Cham (Eastern)	Austronesian	Marc Brunell
Cherokee	Iroquoian	Durbin Feeling
Cheyenne	Algic	Wayne Leman
Chickasaw	Muskogean	Matthew Gordon
Chinese (Mandarin)	Sino-Tibetan	Li Renzhi
Chinese (Min Nan)	Sino-Tibetan	Cheng-Fu Chen
Chiquitano	Macro-Ge	Harry de Haan
Cholón	Hibito-Cholon	Astrid Alexander-Bakkerus
Culina/Kulina	Arauan	Stefan Dienst
Daba	Afro-Asiatic	Ruth Lienhard
Dan (Gweetawu)	Niger-Congo	Alphonse Mognan Kessegebeu
Dime	Afro-Asiatic	Seyoum G. Mulugeta
Duna	Trans-New Guinea	Lila San Roque
Ese Ejja/Essejja (Peru, Baawaja Kuiñaje)	Tacanan	María C. Chavarría Mendoza
Estonian	Uralic	Petar Kehayov
Eton	Niger-Congo	Mark Van de Velde
Even	Altaic	Aleksey A. Burykin
Evenki (Eastern, Tommot)	Altaic	Tamara Ye. Andreeva
Fijian	Austronesian	Apolonia Tamata
Finnish	Uralic	Jouni Rostila

Language	Family (Ethnologue)	Person(s)
Fon	Niger-Congo	Maxime Adjanohoun, Ines Fiedler
Georgian	Kartvelian	Nino Amiridze, Tinatin Bolkvadze
Greek (Ancient)	Indo-European	Marc Huys, Silvia Luraghi
Greek (Modern)	Indo-European	Eleni Valma
Greenlandic (East)	Eskimo-Aleut	Philippe Mennezier
Greenlandic (West)	Eskimo-Aleut	Michael Fortescue
Guaraní	Tupi	Maura Velázquez-Castillo
Hadza/Bali	Khoisan	Bonny Sands
Hassaniyya	Afro-Asiatic	Catherine Taine-Cheikh
Hausa (Standard & Central (Katsinanci))	Afro-Asiatic	Mahamane L. Abdoulaye, Paul Newman
Hebrew (Ancient)	Afro-Asiatic	David Kummerow, Gary A. Rendsburg
Hebrew (Modern)	Afro-Asiatic	Shlomo Izre'el
Hindi	Indo-European	Lucy Rosenstein, Abhishek Avtans
Huarijio/Warihio	Uto-Aztecan	Rolando Guadalupe Félix Armendáriz
Huichol	Uto-Aztecan	Joseph E. Grimes
Hungarian	Uralic	Marianne Bakró-Nagy, Anna Widmer
Ifugao (Lagawe)	Austronesian	Anne West
Igbo	Niger-Congo	Chinedu Uchechukwu, Victor Manfredi
Ikwere	Niger-Congo	Sylvester Osu
Iraqw	Afro-Asiatic	Maarten Mous
Italian	Indo-European	Graziano Savà
Itelmen	Chukotko-Kamchatkan	Jonathan D. Bobaljik
Iwal	Austronesian	Joel Bradshaw
Japanese	Japanese	Mizuki Miyashita, Nobuko Yoneda
Jarawara/Jaruára	Arauan	Alan Vogel
Jehai	Austro-Asiatic	Niclas Burenhult
Kabardian	North Caucasian	John Colarusso
Kalami	Indo-European	Joan Baart
Kalinga (Lower Tanudan)	Austronesian	Glenn & Jewell Machlan
Kalmyk	Altaic	Ágnes Birtalan
Kammu/Khmu (Yuan)	Austro-Asiatic	Damrong Tayanin
Kanakanabu	Austronesian	Dah-An Ho, Paul Li, Jozef Szakos
Kapampangan	Austronesian	Michael R. M. Pangilinan, Hiroaki Kitano
Karaboro/Kar (Eastern)	Niger-Congo	Klaudia Dombrowsky-Hahn
Karao	Austronesian	Sherri Brainard
Kâte	Trans-New Guinea	Claudia Gerstner-Link
Kaxararí/Kaxarirí	Panoan	Alexandre Couto

Language	Family (Ethnologue)	Person(s)
Kayabí	Tupi	Rose Dobson
Ket	Yenisei	Edward Vajda, Andrey Nefedov
Kewa	Trans-New Guinea	Karl Franklin
Kgalagadi	Niger-Congo	Kems Monaka
Khalkha/Halh Mongolian	Altaic	Ágnes Birtalan
Khanty	Uralic	Anna Widmer
Kilivila/Kiriwina	Austronesian	Gunter Senft
Kilmeri	Trans-New Guinea	Claudia Gerstner-Link
Kissi/Kisi	Niger-Congo	G. Tucker Childs
Kombai	Trans-New Guinea	Lourens de Vries
Konni	Niger-Congo	Mike Cahill
Korafe (Yegha)	Trans-New Guinea	Cindi Farr
Korean	Isolate	Sun-Young Lee, Kiyong Choi, Kyung-Ah Kim, Jae Jung Song, Chungmin Lee
Korowai	Trans-New Guinea	Lourens de Vries
Koryak	Chukotko-Kamchatkan	Valentina R. Dedyk
Kriol (Roper River area)	Creole	Margaret C. Sharpe
Kumak/Nêlêmwa/Nelema	Austronesian	Isabelle Bril
Kwaza/Koaiá	Isolate	Hein van der Voort
Lacandon	Mayan	Suzanne Cook
Lakota	Siouan	Regina Pustet, James Green
Langi	Niger-Congo	Margaret Dunham
Lepcha	Sino-Tibetan	Heleen Plaisier
Libido	Afro-Asiatic	Joachim Crass
Lingala	Niger-Congo	Michael Meeuwis
Luwo	Nilo-Saharan	Joseph Modesto
Machiguenga	Arawakan	Betty A. Snell
Ma'di	Nilo-Saharan	John Mairi Blackings
Malagasy	Austronesian	Charles Randriamasimanana
Malayalam	Dravidian	M.T. Hany Babu, Menon Mythili
Mambay	Niger-Congo	Erik John Anonby
Mandaic/Neo-Mandaic	Afro-Asiatic	Charles Häberl
Mapuche/Mapudungun	Araucanian	Marisa Malvestitti, Antonio Díaz-Fernández Aráoz
Marathi	Indo-European	Kashi Wali
Matengo	Niger-Congo	Nobuko Yoneda
Mauwake	Trans-New Guinea	Liisa & Jouko Berghäll
Mbay	Nilo-Saharan	Madeleine Somte
Menya	Trans-New Guinea	Carl & Pat Whitehead

Language	Family (Ethnologue)	Person(s)
Mixteco (Magdalena Piñasco)	Oto-Manguean	Barbara Hollenbach
Motlav/Mwotlap	Austronesian	Alexandre François
Motu	Austronesian	Andrew James Taylor
Movima	Isolate	Katharina Haude
Muna	Austronesian	René van den Berg
Mutsun/Ohlone (Southern)	Penutian	Lynnika Butler
Naga (Ao)/Ao	Sino-Tibetan	Alec Coupe
Nama (Khoekhoe)	Khoisan	Levi Namaseb, Wilfrid H.G. Haacke
Namia	Sepik-Ramu	Becky Feldpausch
Navajo	Na-Dene	Melvatha R. Chee, Ellavina Perkins, Theodore B. Fernald
Newar	Sino-Tibetan	Kazuyuki Kiryu
Nez Perce	Penutian	Noel Rude
Ngam	Nilo-Saharan	Madeleine Somte
Nganasan (Ust'-Avam)	Uralic	Valentin Yu. Goussev
Nivkh/Gilyak	Isolate	Tohru Kaneko, Itsuji Tangiku
Nubi	Creole	Xavier Luffin
Numbami	Austronesian	Joel Bradshaw
Obolo	Niger-Congo	Enene N. Enene
Ogbah	Niger-Congo	Amadi Ahiamadu
Oneida	Iroquoian	Karin Michelson
Oriya	Indo-European	Manideepa Patnaik
Oromo (West Central, Mecha)	Afro-Asiatic	Tolemariam Fufa
Otomí (Mezquital)	Oto-Manguean	Doris Bartholomew
Pame (Northern)	Oto-Manguean	Scott Berthiaume
Papago-Pima/Tohono	Uto-Aztecan	Madeleine Mathiot
O'odham		
Papiamentu	Creole	Hélène Garrett
Persian	Indo-European	Maziar Toosarvandani
Piemontese	Indo-European	Mauro Tosco
Popoluca (Sierra)	Mixe-Zoque	Salomé Gutiérrez Morales
Purepecha/Tarascan	Isolate	Fernando Nava, Ricardo Maldonado
Puyuma	Austronesian	Stacy Teng
Qawasqar/Kawesqar	Alacalufan	José Pedro Viegas Barras
Rapa Nui	Austronesian	Viki Haoa Cardinali
Rejang	Austronesian	Richard McGinn
Rembarunga/Rembarrnga	Australian	Adam Saulwick
Reyesano	Tacanan	Antoine Guillaume
Rukai (Budai, Kucapungan)	Austronesian	Cheng-Fu Chen

Language	Family (Ethnologue)	Person(s)
Rwanda	Niger-Congo	Alexandre Kimenyi
Saami (Northern)	Uralic	Pekka Sammallahti
Saaroa	Austronesian	Paul Li
Sandawe	Khoisan	Helen Eaton
Sango	Creole	Madeleine Somte
Sateré-Mawé	Tupi	Sérgio Meira
Savosavo	East Papuan	Claudia Wegener
Semelai	Austro-Asiatic	Nicole D. Kruspe
Seri	Hokan	Stephen Marlett
Shanenawa	Panoan	Lincoln Almir Amarante Ribeiro, Gláucia Vieira Cândido
Sheko	Afro-Asiatic	Anne-Christie Hellenthal
Somali (Mogadishu)	Afro-Asiatic	Mauro Tosco
Sorbian (Upper)	Indo-European	Leńka Scholze
Squamish	Salishan	Peter Jacobs
Suabo/Inanwatan	Trans-New Guinea	Lourens de Vries
Subiya/Kuhane/Ciikuhane	Niger-Congo	Ndana Ndana
Suyá	Macro-Ge	Cilene Campetela
Swahili	Niger-Congo	Assibi A. Amidou
Takwane	Niger-Congo	Oliver Kröger, Jeff & Peg Shrum
Tatuyo	Tucanoan	Elsa Gómez-Imbert
Tepehua (Pisaflores)	Totonacan	Jim Watters
Tepehua (Tlachichilco)	Totonacan	Jim Watters
Terêna	Arawakan	Onilda Sanches
Teribe	Chibchan	Juan Diego Quesada
Thai	Tai-Kadai	David Smyth, Pittayawat Pittayaporn
Tharaka/Kitharaka	Niger-Congo	Peter Muriungi
Thayore/Kuuk Thaayorre	Australian	Alice Gaby
Tibetan (Shigatse/Xigazê)	Sino-Tibetan	Felix & Chungda Haller
Tikar	Niger-Congo	Ellen Jackson
Totonac (Filomena Mata-Coahuilán)	Totonacan	Teresa Ann McFarland
Toura	Niger-Congo	Thomas Bearth
Trió	Carib	Sérgio Meira
Tsamai/Ts'amakko	Afro-Asiatic	Graziano Savà
Tsimshian (Coast)/Sm'algyax	Penutian	Margaret S. Anderson
Tswana	Niger-Congo	Andy Chebanne
Tswapong	Niger-Congo	Bennett Maifala
Tukang Besi	Austronesian	Mark Donohue

Language	Family (Ethnologue)	Person(s)
Turkish	Altaic	Ceyhan Temürcü
Udihe	Altaic	Maria V. Tolskaya
Vietnamese	Austro-Asiatic	Thu Thi Anh Nguyen
Warlmanpa	Australian	David Nash
Wichita	Caddoan	David S. Rood
Wobé/Wè (Northern)	Niger-Congo	Inge Egner
Wolof (Dakar)/Waro-Waro	Niger-Congo	Stéphane Robert
Xerênte/Akwê-Xerente	Macro-Ge	Sinval Martins de Sousa Filho
Xhosa	Niger-Congo	Bertie Neethling
Yabem	Austronesian	Joel Bradshaw
Yami	Austronesian	Victoria Der-Hwa Rau, Maa-Neu Dong
Yi (Sichuan/Liangshan)	Sino-Tibetan	Andy & Emily Eatough
Yoruba	Niger-Congo	Akinloye A. Ojo, Victor Manfredi
Yugambal-Bandjalang	Australian	Margaret C. Sharpe
Zande	Niger-Congo	Raymond Boyd
Zapotec (San Bartolomé Loxicha)	Oto-Manguean	Rosemary Beam de Azcona
Zo'é	Tupi	Ana Suelly Arruda Câmara Cabral

Appendix B.1

The questionnaire: the “regular” version

This appendix contains the “regular” version of the questionnaire I used (cf. Section I.7).

Part I. Interrogative pronouns

What are the words for ‘who?’, ‘what?’, ‘which (one)?’ and ‘where?’ in LANGUAGE. Do these words distinguish grammatical gender, have plural forms or have defective case paradigms? Can they function as, or are formally related to indefinite expressions?

Part II. Translate in LANGUAGE

- (1) *What is the baby? A girl or a boy?*
- (2) *What is the seller in the shop/the person in that hut/etc.? A man or a woman?*

Another possible situation:

A: *I joined the army yesterday. Half of the officers are men, the others are women.*

B: *And what is the commander?*

A: *(It’s) a man.*

- (3)
 - a. *What is John/your brother? Is he a smith or a carpenter?* (or any other two typical names of occupation, caste, etc.).
 - b. A: *What is he?* B: *He is British/German.*
 - c. Does an equational construction of the type *What is John?* (i.e., where the interrogative pronoun ‘what?’ is equated to a noun designating a person) make any sense at all in LANGUAGE.
- (4) *What is your name?*
- (5) Is it possible to translate in LANGUAGE something like:
 - a. *Who is your name?*
 - b. *Who are you called?*

(If yes, please provide the translations.)

- (6) How does one ask about the name of an unknown object in LANGUAGE?
For instance, you see an unknown tree:
- a. *What is this (tree)?*
 - b. *What is the name of this tree?*
 - c. *How is this tree called?*
- (If yes, please provide the translations.)
- (7) Is it possible to use the equivalent of ‘who?’ in the translations of the questions in (6)? If yes, is it possible with any object or are there restrictions?
- (8) How does one ask about the name of an unknown animal in LANGUAGE? Possible situations: you are in the zoo and you see an unknown animal/you see a picture of an unknown animal/you catch a glimpse of an animal disappearing into the distance (behind the bush, diving under water, whatever).
- a. *What is this (animal)?*
 - b. *What is the name of this animal?*
 - c. *How is this animal called?*
- (If yes, please provide the translations.)
- (9) Is it possible to use the equivalent of ‘who?’ in the translations of the questions in (8)? If yes, is it possible with any living being or are there restrictions on the use of ‘who?’? (e.g., only for higher animates, for big animals, for animals and birds, but not for snails or dragonflies)
- (10) *What is X* [where X is a name of any object in LANGUAGE] *in English/ French/ etc.*? For instance, an English speaker learning French may ask: *What is “January”/ “Christmas”/ “apple”/ “rabbit”/ “victory”/ etc. in French?*
- (11) Is it possible to use the equivalent of ‘who?’ in the translation of (11)? If yes, is it possible with any kind object or are there restrictions? (e.g., only for the names of months, holidays, animals, etc.)
- (12) How does one ask about the name of an unknown place in LANGUAGE?
- a. *What’s the name of the place here?*
 - b. *How is this place called?*
- (Possible answers: *It’s Brooklyn/ the Grand Canyon/ Twin Swamps/ etc.*)
- (13) Is it possible to use the equivalent of ‘who?’ in the translation of (12)? If

yes, is it possible with any places or are there restrictions? (e.g., only for the names of settlements)

- (14) Someone walks into a room with a swelling on his arm, a wound or whatever, do you say in LANGUAGE: *Who bit/ stung you?* or *What bit/ stung you?*
- (15) The situation: you know that Ann and Mary are somehow related to each other, but you do not know the exact nature of this relation.
 You: *What/Who is Mary to you?*
 Ann: *She is my sister-in-law.*
- (16) The situation: you know that Mary is Ann’s sister-in-law and you would like to know how good their relation is.
 You: *What/Who is Mary to you?*
 Ann: *She is a good friend of mine.*
- (17) A: *What/Who is Stalin/George W. Bush/God/etc. to you?* (any well-known person/deity/etc. with whom the person B definitely does not have any kind of kinship relation)
 B: *He is ...* (anything you would find to be a possible answer)
- (18) Is it possible to use the word for ‘what?’ in LANGUAGE attributively with a noun denoting a person? Something like *what man?* which would, for instance, mean ‘what kind of man?’ or ‘which man?’?
- (19) Is it possible to use the word for ‘who?’ in LANGUAGE attributively with a noun denoting a person? Something like *who man?* which would, for instance, mean ‘which man?’ or ‘what kind of man?’?
- (20) Is it possible to use the word for ‘what?’ in LANGUAGE attributively with a noun denoting an object/an animal/a place/etc.? Something like *what knife?* which would, for instance, mean ‘what kind of knife?’ or ‘which knife?’?
- (21) Is it possible to use the word for ‘who?’ in LANGUAGE attributively with a noun denoting an object/an animal/a place/etc.? Something like *who knife?* which would, for instance, mean ‘which knife?’ or ‘what kind of knife?’?

Appendix B.2

The questionnaire: the “light” version

This appendix contains the “light” version of the questionnaire I used (cf. Section I.7).

Part I. Three general questions

- (1) What are the words for ‘who?’, ‘what?’, ‘which (one)?’ and ‘where?’ in LANGUAGE. Do these words distinguish grammatical gender, have plural forms or have defective case paradigms? Can they function as, or are formally related to indefinite expressions?
- (2) Does an equational construction of the English type ‘What is John?’ (i.e., where the interrogative pro-noun ‘what?’ is equated to a noun designating a person) make any sense at all in LANGUAGE? (various possible meanings which this construction may have can be found in the Questionnaire)
- (3) Does an equational construction of the type ‘Who is *X*?’ (where *X* is a noun that does not designate a person, by preference it designates a thing) make any sense at all in LANGUAGE? (various possible meanings which this construction may have can be found in the Questionnaire, e.g. in Indonesian one would ask something like ‘Who is your name?’)

Part II. Four additional questions

- (1) Is it possible to use the word for ‘what?’ in LANGUAGE attributively with a noun denoting a person? Something like *what man?* which would, for instance, mean ‘what kind of man?’ or ‘which man?’?
- (2) Is it possible to use the word for ‘who?’ in LANGUAGE attributively with a noun denoting a person? Something like *who man?* which would, for instance, mean ‘which man?’ or ‘what kind of man?’?
- (3) Is it possible to use the word for ‘what?’ in LANGUAGE attributively with a noun denoting an object/an animal/a place/etc.? Something like *what knife?* which would, for instance, mean ‘what kind of knife?’ or ‘which knife?’?
- (4) Is it possible to use the word for ‘who?’ in LANGUAGE attributively with a noun denoting an object/an animal/a place/etc.? Something like *who knife?* which would, for instance, mean ‘which knife?’ or ‘what kind of knife?’?

Appendix C

The sample (cf. Section I.6)

Language	Family (Ethnologue)	Language	Family (Ethnologue)
Aari	Afro-Asiatic	Amanab	Trans-New Guinea
Abaza (Tapanta)	North Caucasian	Amarakaeri/Harakmbet	Harakmbet
Abidji	Niger-Congo	Ambae (West)/Opa	Austronesian
Abkhaz	North Caucasian	Ambai	Austronesian
Abui	Trans-New Guinea	Ambonese (Malay)	Creole
Abujmaria	Dravidian	Amele	Trans-New Guinea
Abun	West Papuan	Amharic	Afro-Asiatic
Abung/Lampung Nyo (Melinting)	Austronesian	Amis (Central)	Austronesian
Aceh	Austronesian	Amo/Ba	Niger-Congo
Achagua	Arawakan	Amuesha	Arawakan
Achi' (Cubulco)	Mayan	Andamanese (Great)	Andamanese
Acholi	Nilo-Saharan	Andoque	Isolate
Achuar	Jivaroan	Aneityum	Austronesian
Acoma/Keres (Western)	Keres	Anindilyakwa	Australian
Adynyamathanha	Australian	Anuak/Anywa	Nilo-Saharan
Afar/Qafar	Afro-Asiatic	Apalai	Carib
Aguaruna	Jivaroan	Apali	Trans-New Guinea
Ainu	Isolate	Apatani	Sino-Tibetan
Ajyíninka (Apurucayali)	Arawakan	Apinayé	Macro-Ge
Akan	Niger-Congo	Apurinã	Arawakan
Akateka (San Miguel Acatán)	Mayan	Arabana	Australian
Akawaio/Ingarikó	Carib	Arabela	Zaparoan
Akhvakh (Northern)	North Caucasian	Arabic (Chadian)	Afro-Asiatic
Akkadian	Afro-Asiatic	Arabic (Egyptian)	Afro-Asiatic
Alabama	Muskogean	Arabic (Judeo-Iraqi, Baghdadi)	Afro-Asiatic
Alak	Austro-Asiatic	Arabic (Mesopotamian)	Afro-Asiatic
Alamblak	Sepik-Ramu	Arabic (Moroccan)	Afro-Asiatic
Alawa	Australian	Arabic (North Levantine Spoken, Damascus)	Afro-Asiatic
Albanian	Indo-European	Arabic (Sanaani Spoken)	Afro-Asiatic
Aleut (Western)	Eskimo-Aleut	Arabic (South Levantine Spoken, Negev Bedouin)	Afro-Asiatic
Alutor (Proper)	Chukotko- Kamchatkan	Arabic (Standard)	Afro-Asiatic
Amami-Oshima (Northern, Naze)	Japanese	Araona	Tacanan
		Arapaho	Algic

Language	Family (Ethnologue)	Language	Family (Ethnologue)
Arapesh	Torricelli	Baniva/Baniwa of Guainia	Arawakan
Arawak/Lokono	Arawakan	Baniwa (of Içana)	Arawakan
Archi	North Caucasian	Banoni	Austronesian
Arhuaco/Ika	Chibchan	Barasana	Tucanoan
Armenian (Eastern)	Indo-European	Bardi	Australian
Arrernte (Eastern: Mparntwe Arrernte)	Australian	Baré	Arawakan
Asháninka	Arawakan	Bari	Chibchan
Ashéninka (Perené)	Arawakan	Baruga (Tafota)	Trans-New Guinea
Ashéninka (Pichis)	Arawakan	Basaa (Mbene)	Niger-Congo
Ashéninka (Ucayali-Yurúa)	Arawakan	Bashkir	Altaic
Asmat (Central)	Trans-New Guinea	Basque	Isolate
Assamese	Indo-European	Batak (Karo)	Austronesian
Atayal (Mayrinax)	Austronesian	Batak (Toba)	Austronesian
Atayal (Wulai)	Austronesian	Batek/Bateg Deq	Austro-Asiatic
Athpariya/Athpare	Sino-Tibetan	Batek/Bateg Nong	Austro-Asiatic
Au	Torricelli	Bathari	Afro-Asiatic
Avar	North Caucasian	Baure	Arawakan
Awa Pit	Barbacoan	Bawm	Sino-Tibetan
Awakateka (Aguacatán)	Mayan	Bayali/Dharumbal	Australian
Awara	Trans New-Guinea	Beja/Bedawi (Hadendowa)	Afro-Asiatic
Awetí	Tupi	Bengali	Indo-European
Awngi	Afro-Asiatic	Berber ("continuous" Zenati, Aït Seghrouchen of Oum Jeniba)	Afro-Asiatic
Awtuw	Sepik-Ramu	Berber (Ahaggar Tuareg)	Afro-Asiatic
Aymara	Aymaran	Berber (East Zenati, Tamezret)	Afro-Asiatic
Azerbaijani (North)	Altaic	Berber (Ghadamès)	Afro-Asiatic
Babatana (Babatana)	Austronesian	Berber (Kabyle)	Afro-Asiatic
Babatana (Sisingga)	Austronesian	Berber (Malian Tuareg)	Afro-Asiatic
Badaga	Dravidian	Berber (Mauritanian)/ Zenaga	Afro-Asiatic
Badyara/Badiaranke	Niger-Congo	Berber (Tachelhit)	Afro-Asiatic
Bafanji	Niger-Congo	Berber (Tamazight, Ayt Ndhir)	Afro-Asiatic
Bagvalal	North Caucasian	Berber (Zenati of "oases", Figuig)	Afro-Asiatic
Bahnar (Pleiku)	Austro-Asiatic	Berbice Dutch Creole	Creole
Balangao (Natonin)	Austronesian	Bezhta	North Caucasian
Balinese	Austronesian	Biak	Austronesian
Bambara	Niger-Congo		
Bamu (Pirupiru)	Trans-New Guinea		
Bamu (Sisime)	Trans-New Guinea		
Banda-Linda	Niger-Congo		
Banggarla/Parnkalla	Australian		

Language	Family (Ethnologue)	Language	Family (Ethnologue)
Bicolano (Central, Legaspi)	Austronesian	Candoshi	Isolate
Bidyogo/Bijogo (Kagbagaa)	Niger-Congo	Canela-Kraho/Krahô	Macro-Ge
Bilen/Bilin/Agaw	Afro-Asiatic	Canichana/Kanichana	Isolate
Birhor	Austro-Asiatic	Capanahua	Panoan
Biri	Australian	Car Nicobarese	Austro-Asiatic
Bisa	Niger-Congo	Carapana	Tucanoan
Bislama	Creole	Carib (Island, Dominica)	Arawakan
Blackfoot	Algic	Carib/Galibi	Carib
Blang (Kontoi)	Austro-Asiatic	Carijona	Carib
Bobo Fing/Bobo Madaré	Niger-Congo	Cashinahua	Panoan
Bodo Gadaba/Gutob	Austro-Asiatic	Cavineña	Tacanan
Bolia	Niger-Congo	Cayapa/Chachi	Barbacoan
Bolyu/Palyu	Austro-Asiatic	Cayuvava/Cayubaba	Isolate
Bondo/Remo	Austro-Asiatic	Cebuano	Austronesian
Bontok (Central, Guinaang)	Austronesian	Chácobo	Panoan
Borôro	Macro-Ge	Cham (Eastern)	Austronesian
Boruca	Chibchan	Chamorro	Austronesian
Botlikh	North Caucasian	Chechen	North Caucasian
Brahui	Dravidian	Cherokee	Iroquoian
Brao	Austro-Asiatic	Chewong	Austro-Asiatic
Breton	Indo-European	Cheyenne	Algic
Brokskat	Indo-European	Chichimeca-Jonaz	Oto-Manguean
Bru (?Eastern)	Austro-Asiatic	Chickasaw	Muskogean
Budukh	North Caucasian	Chinantec (Comaltepec)	Oto-Manguean
Bugis	Austronesian	Chinantec (Lealao)	Oto-Manguean
Bukiyip	Torricelli	Chinantec (Tepetotutla)	Oto-Manguean
Bulgarian	Indo-European	Chinese (Mandarin)	Sino-Tibetan
Bunaba/Bunuba	Australian	Chinese (Min Nan)	Sino-Tibetan
Bunu (Bu-Nao)	Hmong-Mien	Chinese (Yue)/Cantonese	Sino-Tibetan
Burarra	Australian	Chipaya/Puquina	Uru-Chipaya
Buriat (Russia)	Altaic	Chippewa (Southwestern	Algic
Burji	Afro-Asiatic	Ojibwa, Ponemah)	
Burmese	Sino-Tibetan	Chiquitano	Macro-Ge
Burushaski	Isolate	Chitimacha	Gulf
Buwandik/Bunganditj	Australian	Chol (Tila)	Mayan
Buyang	Tai-Kadai	Chol (Tumbalá)	Mayan
Cabiyari	Arawakan	Cholón	Hibito-Cholon
Cahuilla	Uto-Aztecan	Chong	Austro-Asiatic
Cambodian/Khmer	Austro-Asiatic	Chontal (Tabasco)	Mayan
Camling	Sino-Tibetan	Ch'orti' (Jocotán)	Mayan

Language	Family (Ethnologue)	Language	Family (Ethnologue)
Chrau	Austro-Asiatic	Djeebbana/Ndjébbana	Australian
Chuj (San Mateo Ixtatán)	Mayan	Djiwarli	Australian
Chuj (San Sebastián Coatán)	Mayan	Dogon (Tommo So & Jamsay)	Niger-Congo
Chukchi	Chukotko-Kamchatkan	Dong	Tai-Kadai
Chut/Ruc	Austro-Asiatic	Dumbea	Austronesian
Cinta Larga	Tupi	Dumi	Sino-Tibetan
Cogui/Kogui	Chibchan	Duna	Trans-New Guinea
Colorado	Barbacoan	Dutch	Indo-European
Comanche	Uto-Aztecan	Dyirbal	Australian
Coos (Hanis)	Penutian	Dyugun	Australian
Coptic	Afro-Asiatic	Egyptian (pre-Coptic)	Afro-Asiatic
Cree (Woods: Ile-à-la-Crosse, Algic Cold Lake, La Loche)		Emberá (Northern, Napipí River)	Choco
Cubeo	Tucanoan	Emberá-Saija/Epéna Pedée	Choco
Cuiba (Maibén)	Guahiban	Engenni	Niger-Congo
Culina/Kulína	Arauan	English (Modern)	Indo-European
Cupeño	Uto-Aztecan	English (Old)	Indo-European
Curripaco	Arawakan	Ese Ejja/Essejja (Bolivia)	Tacanan
Daasanach/Dasenech/Geleba	Afro-Asiatic	Ese Ejja/Essejja (Peru, Baawaja Kuiñaje of Tambopata river)	Tacanan
Daba	Afro-Asiatic	Estonian	Uralic
Dagaare	Niger-Congo	Eton	Niger-Congo
Dagik/Masakin	Niger-Congo	Even	Altaic
Dahalik	Afro-Asiatic	Evenki (Eastern, Tommot)	Altaic
Damar (West)	Austronesian	Evenki (Western, Poligus)	Altaic
Dan (Gweetawu)	Niger-Congo	Evenki (Western, Vanavara)	Altaic
Danau	Austro-Asiatic	Ewe	Niger-Congo
Dâw	Maku	Ewondo	Niger-Congo
Dehu/Drehu	Austronesian	Fijian	Austronesian
Dení	Arauan	Finnish	Uralic
Desano	Tucanoan	Fon	Niger-Congo
Dhao/Ndao	Austronesian	French	Indo-European
Dhuwal/Djapu	Australian	Fula/Pulaar	Niger-Congo
Dido/Tsez (Kidero)	North Caucasian	Fulniô/Yatê	Macro-Ge
Diegueño/Kumiai /Jamul	Hokan	Futunan (East)	Austronesian
Tiipay		Fyam/Fyem	Niger-Congo
Dime	Afro-Asiatic	G(a)ta'	Austro-Asiatic
Djaabugay/Dyaabugay	Australian	G(h)odoberi	North Caucasian
Djamindjung/Jaminjung	Australian		
Djawi	Australian		

Language	Family (Ethnologue)	Language	Family (Ethnologue)
Gadaba	Dravidian	Gurinji/Gurindji	Australian
Gagadu/Gaagudju	Australian	Guro	Niger-Congo
Garawa/Garrwa (Eastern)	Australian	Guyani/Kuyani	Australian
Garawa/Garrwa (Western)	Australian	Gwari/Gbagyi	Niger-Congo
Garifuna/Central American	Arawakan	Hadiyya	Afro-Asiatic
Carib/Black Carib (Hopkins & L�ivingston)		Hadza/Bali	Khoisan
Garos	Sino-Tibetan	Halkomelem (Upriver)	Salishan
Gavi�o do Jiparan�/Gavi�o do Rond�nia/Dig�t/Ik�ro	Tupi	Harari	Afro-Asiatic
Gawwada/Dullay	Afro-Asiatic	Hassaniyya	Afro-Asiatic
Gedeo/Darasa	Afro-Asiatic	Hatam	West Papuan
Geez	Afro-Asiatic	Hatam (Mansim/Borai)	West Papuan
Georgian	Kartvelian	Hausa (Standard & Central (Katsinanci))	Afro-Asiatic
German	Indo-European	Hawaiian	Austronesian
Gikuyu	Niger-Congo	Haya	Niger-Congo
Gonja	Niger-Congo	Hdi	Afro-Asiatic
Gooniyandi	Australian	Hebrew (Ancient)	Afro-Asiatic
Grebo	Niger-Congo	Hebrew (Modern)	Afro-Asiatic
Greek (Ancient)	Indo-European	Hindi	Indo-European
Greek (Modern)	Indo-European	Hinukh/Ginukh	North Caucasian
Greenlandic (East)	Eskimo-Aleut	Hmong Daw/White Hmong	Hmong-Mien
Greenlandic (West)	Eskimo-Aleut	Ho	Austro-Asiatic
Guahibo/Sikuani	Guahiban	Huarijio/Warih�o	Uto-Aztecan
Guambiano	Barbacoan	Huastec (San Lu�s Potos�, Tancanhuitz)	Mayan
Guaran�	Tupi	Huave (San Mateo del Mar)	Huavean
Guarekena/Warekena	Arawakan	Huichol	Uto-Aztecan
Guat�	Macro-Ge	Hungarian	Uralic
Guayabero	Guahiban	Hunzib	North Caucasian
Guaymi	Chibchan	Hupd�/Hupd�	Maku
Gudanji/Ngarnga/Ngarnji	Australian	Ibatan (Itbayaten)	Austronesian
Guguyimidjir	Australian	Idoma	Niger-Congo
Gujarati	Indo-European	Ifugao (Bayninan)	Austronesian
Gujari	Indo-European	Ifugao (Lagawe)	Austronesian
Gumbainggir/Kumbaingar	Australian	Igbo	Niger-Congo
Gumuz	Nilo-Saharan	Ignaciano	Arawakan
Gunwinggu/Mayali	Australian	Ijo/Izon/Kolokuma	Niger-Congo
Gunya	Australian	Ik	Nilo-Saharan
Guragone/Gurr-goni	Australian	Ikwere	Niger-Congo
Gureng Gureng	Australian	Imonda	Trans-New Guinea

Language	Family (Ethnologue)	Language	Family (Ethnologue)
Iñapari	Arawakan	Kalami	Indo-European
Indonesian	Austronesian	Kalarko/Kalaaku/Malba	Australian
Inga	Quechuan	Kalenjin/Nandi	Nilo-Saharan
Ingush	North Caucasian	Kalinga (Guinaang)	Austronesian
Iquito	Zaparoan	Kalinga (Limos) (1)	Austronesian
Iraqw	Afro-Asiatic	Kalinga (Limos) (2)	Austronesian
Irarutu/Kasira	Austronesian	Kalinga (Lower Tanudan)	Austronesian
Irish	Indo-European	Kalinga (Upper Tanudan)	Austronesian
Isnag (Dibagat)	Austronesian	Kalkutung/Kalkatungu	Australian
Italian	Indo-European	Kallahan (Kayapa Proper)	Austronesian
Itelmen	Chukotko-Kamchatkan	Kallahan (Keley-i)	Austronesian
Itene/Moré	Chapacura-Wanham	Kalmyk	Altaic
Iu Mien	Hmong-Mien	Kambaata	Afro-Asiatic
Iwal	Austronesian	Kambera	Austronesian
Ixil (Chajul)	Mayan	Kammu/Khmu (Yuan)	Austro-Asiatic
Ixil (Nebaj)	Mayan	Kanakanabu (1)	Austronesian
Jabirrijabirr/Dyaberdyaber	Australian	Kanakanabu (2)	Austronesian
Jah Hut	Austro-Asiatic	Kankanay (Northern)	Austronesian
Jakalteka (Western)	Mayan	Kannada	Dravidian
Japanese	Japanese	Kanoê/Kapishana	Isolate
Jarawara/Jaruára	Arauan	Kapampangan	Austronesian
Jaru/Djaru	Australian	Kaqchikel (most dialects)	Mayan
Javanese	Austronesian	Karaboro/Kar (Eastern)	Niger-Congo
Jawoyn/Djauan	Australian	Karajá	Macro-Ge
Jehai	Austro-Asiatic	Karajarri/Karadjeri	Australian
Jeng/Cheng	Austro-Asiatic	Karao	Austronesian
Jingulu/Djingili	Australian	Karata	North Caucasian
Jola-Fonyi/Diola-Fogny	Niger-Congo	Karitiâna	Tupi
Jowulu	Niger-Congo	Karlamay	Australian
Ju/'hoan	Khoisan	Karok	Hokan
Juang	Austro-Asiatic	Kashmiri	Indo-European
Jula	Niger-Congo	Kâte	Trans-New Guinea
Kab(b)a (of Paoua)	Nilo-Saharan	Katu (Eastern)	Austro-Asiatic
Kabardian	North Caucasian	Kaurna	Australian
Kadiwéu	Mataco-Guaicuru	Kawaiisu	Uto-Aztecan
Kaingáng	Macro-Ge	Kaxarari/Kaxariri	Panoan
Kairiru	Austronesian	Kayabi	Tupi
Kaixana	Arawakan	Kayah (Eastern)/Red Karen	Sino-Tibetan
Kala Lagaw Ya	Australian	Kayapó/Mebengokre	Macro-Ge
		Kayardild	Australian

Language	Family (Ethnologue)	Language	Family (Ethnologue)
Kei (Tanimbar)	Austronesian	Korowai	Trans-New Guinea
Kensiu	Austro-Asiatic	Korwa/Ernga/Singli	Austro-Asiatic
Kenuzi-Dongola/Dongolawi	Nilo-Saharan	Koryak	Chukotko-
Kera	Afro-Asiatic		Kamchatkan
Ket	Yenisei	Kota	Dravidian
Kewa	Trans-New Guinea	Kpelle	Niger-Congo
Kgalagadi	Niger-Congo	Krenak/Botocudo/Borun	Macro-Ge
Khalkha/Halh Mongolian	Altaic	Kriol (Roper River area)	Creole
Khana/Kana	Niger-Congo	Krongo	Niger-Congo
Khanty	Uralic	Kryts	North Caucasian
Kharia	Austro-Asiatic	Kui	Dravidian
Khasi	Austro-Asiatic	Kuku-Yalanji	Australian
Khinalug	North Caucasian	Kumak/Nêlêmwa	Austronesian
Khvarshi (Inkhokari)	North Caucasian	Kuna/Cuna (Border)	Chibchan
K'iche'/Quiche	Mayan	Kunama	Nilo-Saharan
Kilivila/Kiriwina	Austronesian	Kunimaipa	Trans-New Guinea
Kilmeri	Trans-New Guinea	Kuot	East Papuan
Kintaq (Bong)	Austro-Asiatic	Kurdish (Central)	Indo-European
Kiowa	Kiowa Tanoan	Kurdish (Northern)/	Indo-European
Kiribati	Austronesian	Kurmanji	
Kissi/Kisi	Niger-Congo	Kurux	Dravidian
Kiwai (Southern, Doumori)	Trans-New Guinea	Kusunda	Isolate
Kiwai (Southern, Iasa)	Trans-New Guinea	Kuvi	Dravidian
Kiwai (Southern, Turituri)	Trans-New Guinea	Kuy (Ntua)	Austro-Asiatic
Klamath-Modoc	Penutian	Kwaio	Austronesian
Koasati	Muskogean	Kwamera	Austronesian
Kobon	Trans-New Guinea	Kwaza/Koaiá	Isolate
Kodagu	Dravidian	Kwini	Australian
Kokota	Austronesian	Lacandon	Mayan
Kolami (Kolami & Naiki)	Dravidian	Ladakhi	Sino-Tibetan
Kombai	Trans-New Guinea	Lahu	Sino-Tibetan
Konda	Dravidian	Lak	North Caucasian
Kongo (San Salvador)	Niger-Congo	Lakota	Siouan
Konni	Niger-Congo	Langi	Niger-Congo
Korafe (Yegha)	Trans-New Guinea	Lango	Nilo-Saharan
Korana	Khoisan	Lanoh (Jengjeng)	Austro-Asiatic
Korean	Isolate	Lanoh (Yir)	Austro-Asiatic
Koreguaje	Tucanoan	Laragia/Larrakiya/	Australian
Korku	Austro-Asiatic	Gulumirrgin	
Koromfe	Niger-Congo	Lardil	Australian

Language	Family (Ethnologue)	Language	Family (Ethnologue)
Latin	Indo-European	Maltese	Afro-Asiatic
Latvian	Indo-European	Malto (Malpaharia)	Dravidian
Laven	Austro-Asiatic	Malto (Sauria)	Dravidian
Lavukaleve	East Papuan	Mam (Northern, San Pedro Necta)	Mayan
Lawa (Western)	Austro-Asiatic	Mambay	Niger-Congo
Laz	Kartvelian	Manam	Austronesian
Lele	Afro-Asiatic	Manchu	Altaic
Lenakel	Austronesian	Manda	Dravidian
Lepcha	Sino-Tibetan	Mandahuaca/Mandawaka	Arawakan
Leti	Austronesian	Mandaic/Neo-Mandaic	Afro-Asiatic
Lezgian	North Caucasian	Mangarayi	Australian
Libido	Afro-Asiatic	Mangarla/Mangala	Australian
Lillooet	Salishan	Manikion/Sougb	East Bird's Head
Limilngan	Australian	Manitenére/Mantinera	Arawakan
Lingala	Niger-Congo	Manobo (Agusan)	Austronesian
Lithuanian	Indo-European	Manobo (Western Bukidnon)	Austronesian
Lobi	Niger-Congo	Mansaka	Austronesian
Loniu	Austronesian	Maori	Austronesian
Louisiana Creole	Creole	Mapuche/Mapudungun	Araucanian
Luo	Nilo-Saharan	Mara	Australian
Lurija/Luritja	Australian	Maranunggu/Emmi/Warrgat	Australian
Luvale	Niger-Congo	Marathi	Indo-European
Luwo	Nilo-Saharan	Margany	Australian
Maba	Nilo-Saharan	Marghi	Afro-Asiatic
Macá	Mataco-Guaicuru	Mari (Western/High)	Uralic
Machiguenga	Arawakan	Maricopa	Hokan
Macuna	Tucanoan	Marringarr/Maringarr	Australian
Macushi	Carib	Marrithiyel/Marithiel	Australian
Ma'di	Nilo-Saharan	Martu Wangka (Yulparija)	Australian
Mah Meri/Besisi	Austro-Asiatic	Martuthunira/Martuyhunira	Australian
Mailu/Magi	Trans-New Guinea	Masalit	Nilo-Saharan
Maipure	Arawakan	Mataco (Wichi Lhamtés Nocten)	Mataco-Guaicuru
Maithili	Indo-European	Mataco (Wichi Lhamtés Vezoz)	Mataco-Guaicuru
Makah	Wakashan	Matengo	Niger-Congo
Mal/Thin	Austro-Asiatic	Matis	Panoan
Malagasy	Austronesian	Matsés	Panoan
Malayalam	Dravidian	Maung	Australian
Maldivian/Dhivehi	Indo-European		
Ma(a)le	Afro-Asiatic		
Malecite-Passamaquoddy	Algic		

Language	Family (Ethnologue)	Language	Family (Ethnologue)
Mauwake	Trans-New Guinea	Muinane	Witotoan
Maxakalí	Macro-Ge	Muk-thang/Gaanay/Kurnai	Australian
Maya (Itza')	Mayan	Mullukmulluk/Malak-	Australian
Maya (Mopan, San Luis)	Mayan	Malak/Nguluwongga	
Maya (Yucatán)	Mayan	Muna	Austronesian
Maybrat/Mai Brat	West Papuan	Mundang	Niger-Congo
Mbay	Nilo-Saharan	Mundari	Austro-Asiatic
Mbosi/Mboshi	Niger-Congo	Mundurukú	Tupi
Mbum	Niger-Congo	Murle	Nilo-Saharan
Mehinaku	Arawakan	Mussau-Emira	Austronesian
Mehri/Mahri	Afro-Asiatic	Mutsun/Ohlone (Southern)	Penutian
Meit(h)ei/Manipuri	Sino-Tibetan	Mwaghavul/Mupun	Afro-Asiatic
Mekeo	Austronesian	Nadëb	Maku
Melanau/Milano	Austronesian	Naga (Angami)	Sino-Tibetan
Mende	Niger-Congo	Naga (Ao)/Ao	Sino-Tibetan
Menya	Trans-New Guinea	Naga (Sumi)/Sema	Sino-Tibetan
Meyah	East Bird's Head	Naga (Tangkhul)	Sino-Tibetan
Mikir/Karbi	Sino-Tibetan	Naga Pidgin	Creole
Minriq/Mendriq	Austro-Asiatic	Nahali/Nihali	Isolate
Mintil	Austro-Asiatic	Nahuatl (Tetelcingo)	Uto-Aztecan
Miriwoong/Miriwung	Australian	Nakara/Nagara/Kokori	Australian
Mirning/Mirninny	Australian	Nama (Khoekhoe)	Khoisan
Mixe (Coatlán)	Mixe-Zoque	Nambikuára (Southern)	Nambiquaran
Mixteco (Chalcatongo)	Oto-Manguean	Nambo (Parb)	Trans-New Guinea
Mixteco (Magdalena Piñasco)	Oto-Manguean	Namia	Sepik-Ramu
Miya	Afro-Asiatic	Nanai (Torgon/Naykhin)	Altaic
Miyako	Japanese	Nangikurrunggurr/	Australian
Mocoví	Mataco-Guaicuru	Ngankikurungkurr	
Mokilese	Austronesian	Narrinyeri (Keramin/Kureinji)	Australian
Mon	Austro-Asiatic	Narrinyeri (Ngayawang)	Australian
Mongo-Nkundo	Niger-Congo	Narrinyeri (Yitha-Yitha/	Australian
Montagnais	Algic	Dardi-Dardi)	
Monumbo	Torricelli	Narrinyeri (Yuyu/Ngarrikt)	Australian
Motlav/Mwotlap	Austronesian	Nateni	Niger-Congo
Motu	Austronesian	Navajo	Na-Dene
Movima	Isolate	Ndyuka	Creole
Mpur	West Papuan	Negidal (Nizovskiy)	Altaic
Mudburra	Australian	Nengone	Austronesian
Mufian	Torricelli	Newar	Sino-Tibetan
		Nez Perce	Penutian

Language	Family (Ethnologue)	Language	Family (Ethnologue)
Ngadjunmaya	Australian	Orok	Altaic
Ngalakan	Australian	Oromo (Borana)	Afro-Asiatic
Ngam	Nilo-Saharan	Oromo (West Central, Mecha)	Afro-Asiatic
Nganasan (Ust'-Avam)	Uralic	Otomí (Mezquital)	Oto-Manguean
Ngandi	Australian	Ottawa/Nishnaabemwin	Algic
Ngarinyin/Ngarinjin/	Australian	Paama/Paamese	Austronesian
Ungarinjin		Paccamalh/Patjtjamalh/	Australian
Ngbaka	Niger-Congo	Wadjiginy	
Ngeq	Austro-Asiatic	Pacoh	Austro-Asiatic
Ngiti	Nilo-Saharan	Páez	Paezan
Ngiyambaa	Australian	Paiute Southern (Kaibab)	Uto-Aztecan
Ngombe (Ligenza)	Niger-Congo	Paiwan	Austronesian
Nhanda	Australian	Palauan	Austronesian
Nhirrpi	Australian	Palaung (Pale)	Austro-Asiatic
Nias	Austronesian	Palaung (Shwe)	Austro-Asiatic
Nimanbur(ru)	Australian	Palu'e	Austronesian
Nivkh/Gilyak	Isolate	Pame (Northern)	Oto-Manguean
Nomatsiguenga	Arawakan	Panamint/Timbisha Shoshone/	Uto-Aztecan
Noon	Niger-Congo	Tümpisa Shoshone (Central)	
Ntomba-Inongo	Niger-Congo	Panamint/Timbisha Shoshone/	Uto-Aztecan
Nubi	Creole	Tümpisa Shoshone (Eastern)	
Nukak Makú	Maku	Pangasinan	Austronesian
Numbami	Austronesian	Panjabi/Punjabi	Indo-European
Nung	Sino-Tibetan	Panyjima/Panytyima	Australian
Nung	Tai-Kadai	Papago-Pima/Tohono	Uto-Aztecan
Nungali	Australian	O'odham	
Nunggubuyu	Australian	Papiamentu	Creole
Nyahkur	Austro-Asiatic	Par(a)katêjê/Gavião of Pará	Macro-Ge
Nyangumarta	Australian	Pareci/Paresi/Haliti	Arawakan
Nyikina/Nyigina	Australian	Parenga/Gorum	Austro-Asiatic
Nyulnyul	Australian	Parji/Duruwa	Dravidian
Nyungar/Nyunga	Australian	Paulohi	Austronesian
Obolo	Niger-Congo	Paumari	Arauan
Ofayé/Opayé	Macro-Ge	Pech	Chibchan
Ogbah	Niger-Congo	Pengo	Dravidian
Oirata	Trans-New Guinea	Pero	Afro-Asiatic
Ona/Selknam	Chon	Persian	Indo-European
Oneida	Iroquoian	Piapoco	Arawakan
Oriya	Indo-European	Piaroa	Salivan
Oroch	Altaic	Piemontese	Indo-European

Language	Family (Ethnologue)	Language	Family (Ethnologue)
Pilagá	Mataco-Guaicuru	Rwanda	Niger-Congo
Pima Bajo	Uto-Aztecan	Saami (Northern)	Uralic
Pintupi	Australian	Saaroa	Austronesian
Pipil	Uto-Aztecan	Sabanês	Nambiquaran
Pirahã	Mura	Sabu/Hawu/Sawu	Austronesian
Piro/Yine	Arawakan	Sabüm/Sabum	Austro-Asiatic
Pitjantjatjara	Australian	Saho (Asa-Awurta & Tarua)	Afro-Asiatic
Pitta Pitta	Australian	Saho (Irob & Minifero)	Afro-Asiatic
Pohnpeian	Austronesian	Saisiyat (Tungho)	Austronesian
Polish	Indo-European	Sakao	Austronesian
Pomo (Southeastern)	Hokan	Sakapulteka (Sacapulas Centro)	Mayan
Popoluca (Sierra)	Mixe-Zoque	Sakirabiá/Mekens Sakirabiat	Tupi
Poqom (Poqomam, Poqomchi')	Mayan	Sáliba	Salivan
Portuguese	Indo-European	Sambal (Tinà)	Austronesian
Purepecha/Tarascan	Isolate	Samo (Southern)	Niger-Congo
Purik/Purki	Sino-Tibetan	Samoa	Austronesian
Puyuma	Austronesian	Samtao	Austro-Asiatic
Q'anjob'al (Eastern)	Mayan	Sandawe	Khoisan
Qawasqar/Kawesqar	Alacalufan	Sangir	Austronesian
Q'eqchi'/Kekchi	Mayan	Sango	Creole
Quechua (Ancash)	Quechuan	Santa Ana/Owa	Austronesian
Quechua (Huellaga)	Quechuan	Santali	Austro-Asiatic
Quechua (Imbabura)	Quechuan	Sanumá	Yanomam
Quechua (North Junín)	Quechuan	Sapuan	Austro-Asiatic
Rama	Chibchan	Sateré-Mawé	Tupi
Rapa Nui	Austronesian	Savosavo	East Papuan
Rawa	Trans-New Guinea	Saweru	Geelvink Bay
Rejang	Austronesian	Secoya	Tucanoan
Rembarunga/Rembarrnga	Australian	Seke (Tangbe)	Sino-Tibetan
Resígaro	Arawakan	Semai	Austro-Asiatic
Reyesano	Tacanan	Semaq Beri	Austro-Asiatic
Rikbaktsa	Macro-Ge	Semelai	Austro-Asiatic
Roma	Austronesian	Semnam	Austro-Asiatic
Romani (Vlax/Vlach)	Indo-European	Seneca	Iroquoian
Rukai (Budai, Kucapungan)	Austronesian	Seri	Hokan
Rukai (Labuan)	Austronesian	Shanenawa	Panoan
Rumanian/Romanian	Indo-European	Shehri/Jibbali	Afro-Asiatic
Rundi	Niger-Congo	Sheko	Afro-Asiatic
Russian	Indo-European	Shilluk	Nilo-Saharan

Language	Family (Ethnologue)	Language	Family (Ethnologue)
Shipibo-Conibo	Panoan	Tacana	Tacanan
Shona	Niger-Congo	Tagalog	Austronesian
Shoshoni (Western)	Uto-Aztecan	Tahitian	Austronesian
Sidamo	Afro-Asiatic	Tainae/Ivori	Trans-New Guinea
Silt'e/Silte/East Gurage	Afro-Asiatic	Takelma	Penutian
Sinaugoro	Austronesian	Takwane	Niger-Congo
Siona	Tucanoan	Taliwang	Austronesian
Sipakapense	Mayan	Tamil	Dravidian
Siriono (Yuqui/Jorá)	Tupi	Tapieté	Tupi
Slave	Na-Dene	Tatar	Altaic
Sobei	Austronesian	Tatuyo	Tucanoan
Somali (Mogadishu)	Afro-Asiatic	Tauya	Trans-New Guinea
Somali (Northern)	Afro-Asiatic	Tegali (Rashad/Kom/ Ngakom/Kome)	Niger-Congo
Songhay (Koyra Chiini/ Koyraboro Senni)	Nilo-Saharan	Tektiteka (Tectitán)	Mayan
Soninke	Niger-Congo	Telugu	Dravidian
Soo	Nilo-Saharan	Temiar	Austro-Asiatic
Soqotri	Afro-Asiatic	Teop	Austronesian
Sora/Savara	Austro-Asiatic	Tepehua (Pisaflores)	Totonacan
Sorbian (Upper)	Indo-European	Tepehua (Tlachichilco)	Totonacan
Sotho (Northern)	Niger-Congo	Tepehuan (Southeastern)	Uto-Aztecan
Sotho (Southern)	Niger-Congo	Terêna	Arawakan
Souei	Austro-Asiatic	Teribe	Chibchan
Spanish	Indo-European	Tetela (some dialects)	Niger-Congo
Squamish	Salishan	Tetun	Austronesian
Stieng (Budeh?)	Austro-Asiatic	Thagungwurrung	Australian
Stieng (Bulo)	Austro-Asiatic	Thai	Tai-Kadai
Suabo/Inanwatan	Trans-New Guinea	Tharaka/Kitharaka	Niger-Congo
Suau	Austronesian	Thayore/Kuuk Thaayorre	Australian
Subiya/Kuhane/Ciikuhane	Niger-Congo	Themne	Niger-Congo
Suena	Trans-New Guinea	Tibetan (Shigatse/Xigazê)	Sino-Tibetan
Supyire	Niger-Congo	Ticuna	Isolate
Suruí (Paiter)	Tupi	Tigre	Afro-Asiatic
Susu	Niger-Congo	Tigrigna/Tigrinya/Tigray	Afro-Asiatic
Suyá	Macro-Ge	Tikar	Niger-Congo
Swahili	Niger-Congo	Timucua	Isolate
Swati	Niger-Congo	Tiri/Tinrin	Austronesian
Swedish	Indo-European	Tiwi	Australian
Sye/Sie/Erromanga	Austronesian	Toba	Mataco-Guaicuru
Taba/Makian (East)	Austronesian	Toda	Dravidian

Language	Family (Ethnologue)	Language	Family (Ethnologue)
Tojolobal (Comitán)	Mayan	Tz'utujil/Tzutujil (Eastern: Santiago Atitlán)	Mayan
Tokelauan	Austronesian	Tz'utujil/Tzutujil (Western: San Juan la Laguna)	Mayan
Tol/Jicaque	Isolate	Ubykh	North Caucasian
Tolomako	Austronesian	Udihe	Altaic
Tongan	Austronesian	Ulch	Altaic
Totonaca (Filomena Mata-Coahuilán)	Totonacan	Ulwa/Sumo-Mayangna	Misumalpan
Totonaca (Northern/Xicotepec de Juárez)	Totonacan	Umbindhamu	Australian
Toura	Niger-Congo	Una	Trans-New Guinea
Trió	Carib	Ura	Austronesian
Trique (Copala)	Oto-Manguean	Urarina	Isolate
Trumaí	Isolate	Urdu	Indo-European
Tsakhur	North Caucasian	Uru	Uru-Chipaya
Tsamai/Ts'amakko	Afro-Asiatic	Urubú-Kaapor	Tupi
Tshangla	Sino-Tibetan	Uru-Eu-Wau-Wau/Amondava	Tupi
Tsimané/Mosetén	Isolate	Usan	Trans-New Guinea
Tsimshian (Coast)/Sm'algyax	Penutian	Uspanteka (San Miguel Uspantán)	Mayan
Tsou (Tfuya)	Austronesian	Uspanteka (Uspantán Centro)	Mayan
Tsova-Tush/Bats	North Caucasian	Ute (Kapúuta & Moǵwáci)	Uto-Aztecán
Tswana	Niger-Congo	Vai	Niger-Congo
Tswapong	Niger-Congo	Vanimó	Sko
Tubarão/Aikaná/Aikanã/Wari/Masaká	Isolate	Vengo	Niger-Congo
Tucano	Tucanoan	Vietnamese	Austro-Asiatic
Tukang Besi	Austronesian	Wa	Austro-Asiatic
Tulu	Dravidian	Waboda/Wabuda	Trans-New Guinea
Tungag	Austronesian	Wagaya/Wakaya (Eastern)	Australian
Tunica	Gulf	Wagaya/Wakaya (Western)	Australian
Tuparí	Tupi	Wagiman/Wageman	Australian
Turkana	Nilo-Saharan	Waiwai/Wai Wai	Carib
Turkish	Altaic	Wakawaka (Duungidjau)	Australian
Tuvaluan	Austronesian	Walapai	Hokan
Tuyuca	Tucanoan	Wallisian/Uvea (East)	Austronesian
Tzeltal	Mayan	Walmajarri	Australian
Tzotzil (an unidentified dialect of San Andrés municipality)	Mayan	Wambaya	Australian
Tzotzil (San Andrés, Bochil, Zinacantán, Huixtán)	Mayan	Wanano/Guanano	Tucanoan
		Wandarang/Warndarang	Australian

Language	Family (Ethnologue)	Language	Family (Ethnologue)
Wanyi	Australian	Yabem	Austronesian
Waorani	Isolate	Yagua	Peba-Yaguan
Wapishana/Wapixana	Arawakan	Yámana/Yahgán	Isolate
Wappo	Yuki	Yami	Austronesian
Warao	Isolate	Yamphu	Sino-Tibetan
Wardaman	Australian	Yankunytjatjara	Australian
Wari/Pakaásnovos	Chapacura-Wanham	Yanyuwa	Australian
Warlmanpa	Australian	Yapese	Austronesian
Warlpiri	Australian	Yaqui	Uto-Aztecan
Warray	Australian	Yawalapiti	Arawakan
Warrgamay	Australian	Yawuru	Australian
Warrrnambool	Australian	Yele	East Papuan
Warrwa	Australian	Yi (Sichuan/Liangshan)	Sino-Tibetan
Warumungu	Australian	Yidiny	Australian
Wathawurrung/Kulin/The	Australian	Yinggarda/Yingkarta	Australian
Western Victorian Language		Yogad	Austronesian
Watjarri/Wajarri	Australian	Yokuts (Wikchamni)	Penutian
Waurá	Arawakan	Yoruba	Niger-Congo
Wayampi	Tupi	Yuchi	Isolate
Wayuu/Guajiro	Arawakan	Yucuna	Arawakan
Welsh	Indo-European	Yugambal-Bandjalang	Australian
Wichita	Caddoan	Yuhup	Maku
Wirangu	Australian	Yukaghir	Isolate
Witoto (Murui/Bue)	Witotoan	Yukpa	Carib
Wiyot	Algic	Yukulta/Ganggalida	Australian
Wobé/Wè (Northern)	Niger-Congo	Yup'ik (Central Siberian, St. Lawrence Island)	Eskimo-Aleut
Woiwurrung	Australian	Yuracare	Isolate
Wolio/Baubau	Austronesian	Yurok	Algic
Wolof (Dakar)/Waro-Waro	Niger-Congo	Zande	Niger-Congo
Worora/Worrorra	Australian	Zapotec (San Bartolomé)	Oto-Manguean
Woun Meu/Waunana	Choco	Loxicha)	
Wuzlam	Afro-Asiatic	Zo'é	Tupi
Xamtanga/Khamtanga/Xamir	Afro-Asiatic	Zoque (Chimalapa)	Mixe-Zoque
Xavánte	Macro-Ge	Zulu	Niger-Congo
Xerénte/Akwẽ-Xerente	Macro-Ge	Zuni	Isolate
Xhosa	Niger-Congo	!Xóǀ	Khoisan
Xinca	Isolate	ᄆHōã	Khoisan
Xipaya	Tupi		
Xoklég/Xokleng	Macro-Ge		

Appendix D

The reduced sample (cf. Section II.5)

The number in the first column identifies the language on the map provided at the end of this list. The (sub)type codes given in the last column are explained in Table II.5:2.

	Language	Family (Ethnologue)	Type
1	Alawa	Australian	1b
2	Albanian	Indo-European	4
3	Alutor (Proper)	Chukotko-Kamchatkan	2b
4	Ambai	Austronesian	4
5	Amharic	Afro-Asiatic	3
6	Apali	Trans-New Guinea	3
7	Arabela	Zaparoan	4
8	Arabic (Chadian)	Afro-Asiatic	3
9	Arabic (Standard)	Afro-Asiatic	4
10	Arapaho	Algic	2b
11	Armenian (Eastern)	Indo-European	4
12	Atayal (Mayrinax)	Austronesian	3
13	Awara	Trans New-Guinea	1b
14	Awetí	Tupi	4
15	Aymara	Aymaran	2b
16	Azerbaijani (North)	Altaic	2b
17	Badaga	Dravidian	4
18	Badyara/Badiaranke	Niger-Congo	4
19	Bambara	Niger-Congo	4
20	Barasana	Tucanoan	1a
21	Basque	Isolate	4
22	Beja/Bedawi (Hadendowa)	Afro-Asiatic	1b
23	Bengali	Indo-European	4
24	Berber (Ahaggar Tuareg)	Afro-Asiatic	2b
25	Bicolano (Central, Legaspi)	Austronesian	4
26	Bisa	Niger-Congo	4
27	Bontok (Central, Guinaang)	Austronesian	4
28	Boruca	Chibchan	4
29	Budukh	North Caucasian	2b
30	Buriat (Russia)	Altaic	1b
31	Burmese	Sino-Tibetan	4

	Language	Family (Ethnologue)	Type
32	Burushaski	Isolate	2b
33	Candoshi	Isolate	1b
34	Canela-Kraho/Krahô	Macro-Ge	2b
35	Capanahua	Panoan	4
36	Cheyenne	Algie	2b
37	Chickasaw	Muskogean	2b
38	Chinese (Mandarin)	Sino-Tibetan	4
39	Chinese (Min Nan)	Sino-Tibetan	4
40	Chiquitano	Macro-Ge	2b
41	Culina/Kulina	Arauan	4
42	Cupeño	Uto-Aztecan	4
43	Daba	Afro-Asiatic	2b
44	Dan (Gweetawu)	Niger-Congo	1b
45	Dime	Afro-Asiatic	3
46	Duna	Trans-New Guinea	3
47	Dutch	Indo-European	4
48	English (Modern)	Indo-European	4
49	Eton	Niger-Congo	3
50	Even	Altaic	3
51	Evenki (Eastern, Tommot)	Altaic	2b
52	Evenki (Western, Poligus)	Altaic	3
53	Fijian	Austronesian	3
54	Finnish	Uralic	4
55	Fon	Niger-Congo	4
56	French	Indo-European	4
57	Georgian	Kartvelian	4
58	German	Indo-European	4
59	Greek (Modern)	Indo-European	4
60	Greenlandic (East)	Eskimo-Aleut	2b
61	Hassaniyya	Afro-Asiatic	4
62	Hausa (Standard & Central (Katsinanci))	Afro-Asiatic	4
63	Hebrew (Ancient)	Afro-Asiatic	4
64	Hebrew (Modern)	Afro-Asiatic	1b
65	Hindi	Indo-European	4
66	Huarijio/Warihío	Uto-Aztecan	4
67	Huichol	Uto-Aztecan	4
68	Hungarian	Uralic	4
69	Igbo	Niger-Congo	4
70	Ik	Nilo-Saharan	3

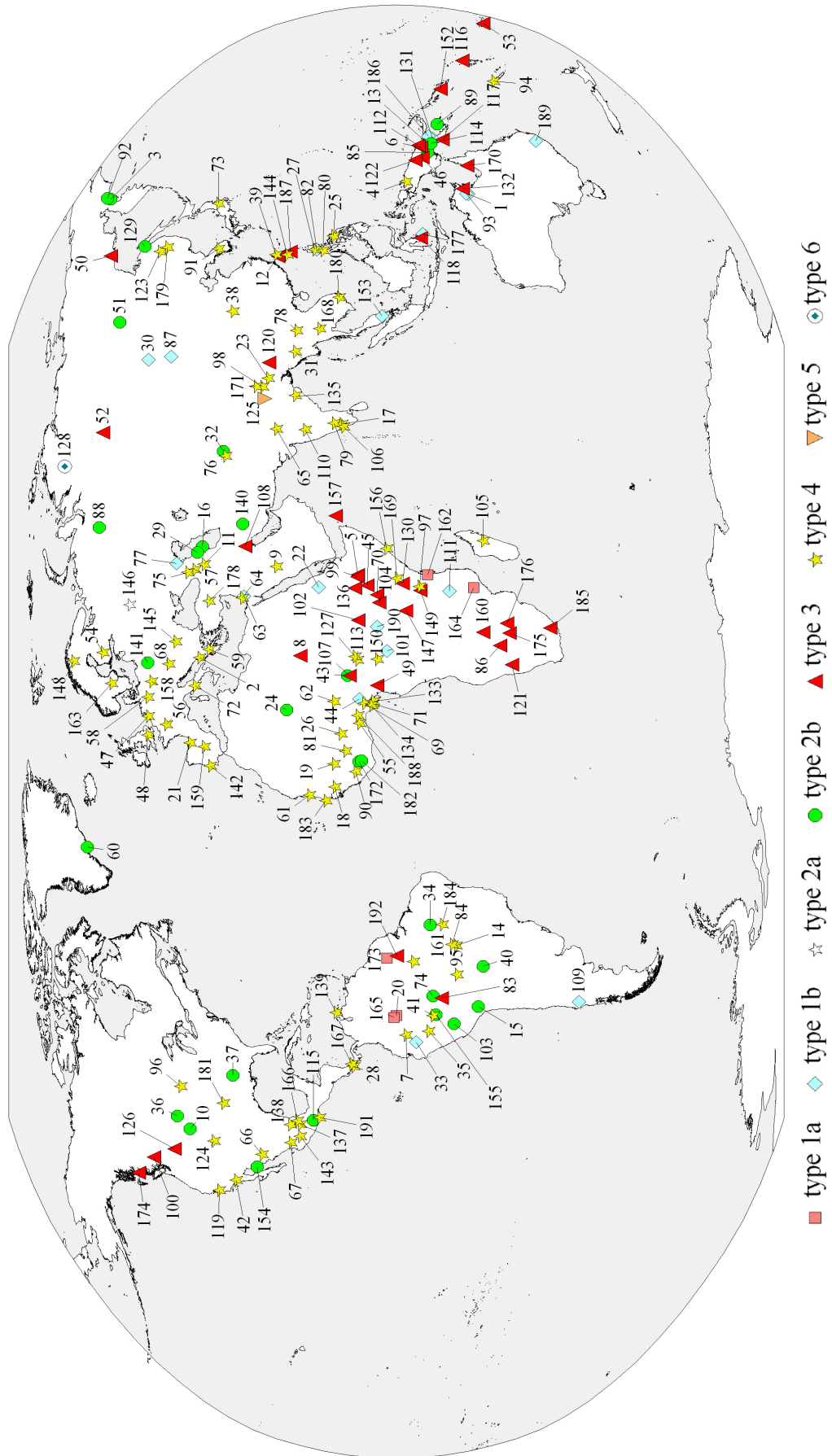
	Language	Family (Ethnologue)	Type
71	Ikwere	Niger-Congo	4
72	Italian	Indo-European	4
73	Japanese	Japanese	4
74	Jarawara/Jaruára	Arauan	2b
75	Kabardian	North Caucasian	4
76	Kalami	Indo-European	4
77	Kalmyk	Altaic	1b
78	Kammu/Khmu (Yuan)	Austro-Asiatic	4
79	Kannada	Dravidian	4
80	Kapampangan	Austronesian	4
81	Karaboro/Kar (Eastern)	Niger-Congo	4
82	Karao	Austronesian	1b
83	Kaxarari/Kaxariri	Panoan	3
84	Kayabi	Tupi	4
85	Kewa	Trans-New Guinea	2b
86	Kgalagadi	Niger-Congo	3
87	Khalkha/Halh Mongolian	Altaic	1b
88	Khanty	Uralic	2b
89	Kilivila/Kiriwina	Austronesian	2b
90	Kissi/Kisi	Niger-Congo	4
91	Korean	Isolate	4
92	Koryak	Chukotko-Kamchatkan	2b
93	Kriol (Roper River area)	Creole	1b
94	Kumak/Nêlêmwa	Austronesian	4
95	Kwaza/Koaiá	Isolate	4
96	Lakota	Siouan	4
97	Langi	Niger-Congo	4
98	Lepcha	Sino-Tibetan	4
99	Libido	Afro-Asiatic	3
100	Lillooet	Salishan	3
101	Lingala	Niger-Congo	1b
102	Luwo	Nilo-Saharan	3
103	Machiguenga	Arawakan	2b
104	Ma'di	Nilo-Saharan	3
105	Malagasy	Austronesian	4
106	Malayalam	Dravidian	4
107	Mambay	Niger-Congo	3
108	Mandaic/Neo-Mandaic	Afro-Asiatic	3
109	Mapuche/Mapudungun	Araucanian	1b
110	Marathi	Indo-European	4

Language	Family (Ethnologue)	Type
111 Matengo	Niger-Congo	1b
112 Mauwake	Trans-New Guinea	3
113 Mbay	Nilo-Saharan	4
114 Menya	Trans-New Guinea	2b
115 Mixteco (Magdalena Piñasco)	Oto-Manguean	2b
116 Motlav/Mwotlap	Austronesian	3
117 Motu	Austronesian	3
118 Muna	Austronesian	3
119 Mutsun/Ohlone (Southern)	Penutian	4
120 Naga (Ao)/Ao	Sino-Tibetan	3
121 Nama (Khoekhoe)	Khoisan	3
122 Namia	Sepik-Ramu	3
123 Nanai (Torgon/Naykhin)	Altaic	4
124 Navajo	Na-Dene	4
125 Newar	Sino-Tibetan	5
126 Nez Perce	Penutian	3
127 Ngam	Nilo-Saharan	4
128 Nganasan (Ust'-Avam)	Uralic	6
129 Nivkh/Gilyak	Isolate	2b
130 Nubi	Creole	3
131 Numbami	Austronesian	1b
132 Nunggubuyu	Australian	3
133 Obolo	Niger-Congo	4
134 Ogbah	Niger-Congo	4
135 Oriya	Indo-European	4
136 Oromo (West Central, Mecha)	Afro-Asiatic	3
137 Otomí (Mezquital)	Oto-Manguean	4
138 Pame (Northern)	Oto-Manguean	4
139 Papiamentu	Creole	4
140 Persian	Indo-European	2b
141 Polish	Indo-European	2b
142 Portuguese	Indo-European	4
143 Purepecha/Tarascan	Isolate	4
144 Rukai (Budai, Kucapungan)	Austronesian	4
145 Rumanian/Romanian	Indo-European	4
146 Russian	Indo-European	2a
147 Rwanda	Niger-Congo	3
148 Saami (Northern)	Uralic	4
149 Sandawe	Khoisan	3
150 Sango	Creole	4

	Language	Family (Ethnologue)	Type
151	Sateré-Mawé	Tupi	4
152	Savosavo	East Papuan	3
153	Semelai	Austro-Asiatic	1b
154	Seri	Hokan	2b
155	Shanenawa	Panoan	2b
156	Somali (Mogadishu)	Afro-Asiatic	4
157	Soqotri	Afro-Asiatic	3
158	Sorbian (Upper)	Indo-European	4
159	Spanish	Indo-European	4
160	Subiya/Kuhane/Ciikuhane	Niger-Congo	3
161	Suyá	Macro-Ge	4
162	Swahili	Niger-Congo	1a
163	Swedish	Indo-European	4
164	Takwane	Niger-Congo	1a
165	Tatuyo	Tucanoan	1a
166	Tepehua (Tlachichilco)	Totonacan	4
167	Teribe	Chibchan	4
168	Thai	Tai-Kadai	4
169	Tharaka/Kitharaka	Niger-Congo	4
170	Thayore/Kuuk Thaayorre	Australian	3
171	Tibetan (Shigatse/Xigazê)	Sino-Tibetan	4
172	Toura	Niger-Congo	2b
173	Trió	Carib	1a
174	Tsimshian (Coast)/Sm'algyax	Penutian	3
175	Tswana	Niger-Congo	3
176	Tswapong	Niger-Congo	3
177	Tukang Besi	Austronesian	1b
178	Turkish	Altaic	4
179	Udihe	Altaic	4
180	Vietnamese	Austro-Asiatic	4
181	Wichita	Caddoan	4
182	Wobé/Wè (Northern)	Niger-Congo	2b
183	Wolof (Dakar)/Waro-Waro	Niger-Congo	4
184	Xerênte/Akwẽ-Xerente	Macro-Ge	4
185	Xhosa	Niger-Congo	3
186	Yabem	Austronesian	1b
187	Yami	Austronesian	3
188	Yoruba	Niger-Congo	4
189	Yugambal-Bandjalang	Australian	1b
190	Zande	Niger-Congo	1b

	Language	Family (Ethnologue)	Type
191	Zapotec (San Bartolomé Loxicha)	Oto-Manguean	4
192	Zo'é	Tupi	3

Map 1. The distribution of the languages of the reduced sample



Appendix E

Dravidian interrogative pronominals (based on Andronov 1978, 1980, 1996; Asher & Kumari 1997; Bhaskarao 1980; Bhat 1989; Burrow & Bhattacharya 1953, 1970; Burrow & Emeneau 1984; Das 1973; Ebert 1996; Emeneau 1955, 1971, 1984; Grierson 1906; Hahn 1911; Krishnamurti & Gwynn 1985; Lehmann 1993; Natarajan 1985; Reddy 1979; Schulze 1911; Shiffman 1999; Pillai & Kothandaraman 1972; Winfield 1928; the papers in the volume edited by Steever 1998c; personal communication from Christiane Pilot-Raichoor)

Language	GS ⁽ⁱ⁾	Set I						Set II	
		SG			PL			HUM	NON<HUM>
		M	F	N	M	F	N		
Southern Dravidian									
Old Tamil	II	ya:vaṇ, evaṇ	ya:vaḷ, evaḷ, ya:r-aḷ	ya:(va)tu, e(va)tu	(y)a:(va)r, evar		ya:(vai), ev(ai)	ya:r	e(:)ṇ, entu, eṇṇa(tu)
Modern Tamil	II	evaṇ evar (HON) ⁽ⁱⁱ⁾	evaḷ	e(:)tu, eḍu, e:ḍi	evanga, eṇṇar, evarkaḷ		ev(ai(kaḷ))	(y)a:r ⁽ⁱⁱⁱ⁾	e(:)ṇ, entu, eṇṇa(tu)
Badaga	II	e:ma, ya:ma	e:va, ya:va	e:du	e:vaka (rare)		e:ve (rare)	da:ra	e:na
Kannada	II	(rare: ya:vanu)	(rare: ya:vaḷu)	ya:vudu	ya:ru		ya:vuvu	= Set I ya:ru ^(iv)	e:nu
Malayalam	VIII	(rare: e(:)van, ya:van)	(rare: e(:)vaḷ, ya:vaḷ)	(rare: e(:)tu, ya:(tu))	(rare: e(:)var, ya:var)		(rare: e(:)va, ya:va)	ya:r, a:r(u)	entə, eṇṇə, entu, ettu ^(v)
Kodagu	II	e:vē, ya:vē	e:va, ya:va	e:di, ya:di	?(evangaḷ)		e:vu, ya:vu	(d)a:rī	enni

Language	GS ⁽ⁱ⁾	Set I						Set II	
		SG			PL			HUM (‘who?’)	NON<HUM> (‘what?’)
		M	F	N	M	F	N		
Kota	II	<i>e(:)vṃ</i>	<i>e(:)v!</i>	(see Set II)	<i>e(:)vr</i>		(see Set II)	<i>(d)a:r, ida:r</i>	<i>e(:)d, e(d)n,</i> <i>eṇ</i>
Tulu	II	(see Set II)		<i>ovu</i>	(see Set II)		<i>oyikulu</i>	<i>e:ru, e:raṇe^(vi)</i>	<i>(da:ne, ja:ne)^(vi)</i>
Toda	VIII	predicative & NOM: <i>e:θ(u)</i> (predicative & very rare & other cases: <i>e:n-</i>)		(predicative & very rare & other cases: <i>e:θ-</i>)	predicative (& NOM?): <i>e:θ(u), e:θum</i> (very rarely: <i>e:θ-a:m</i> (NOM), <i>e:θa:n-</i> (other cases))			<i>o:ry</i> , predica- tive: <i>ory(u)</i>	<i>in, i:</i> ; predi- cative: <i>in(u)</i>
South-Central									
Dravidian									
Old Telugu	V	<i>evaNḍu</i>	<i>ev(v)ate^(vii)</i> <i>e:di, e(d)di, e:yadi^(vii)</i>	<i>e:mi^(vii)</i>	<i>evaru, evva:ralu, ev(v)anḍru</i>	<i>e:vi, ev(v)i,</i> <i>eyyavi</i>		?	?
Modern Telugu	V	<i>ewaḍu</i>	<i>ewate^(viii)</i> <i>e:di, ed(d)i^(viii)</i>		<i>ewaru</i>	<i>e:wi</i>		= Set I <i>ewaru^(viii)</i>	<i>e:m(i), e:miti,</i> <i>enti^(vii)</i>
Muria Gondi	VI	<i>boR</i>	<i>bodd, bad</i>		<i>bor</i>	<i>bovv(es(k))</i>		= Set I <i>bor^(ix)</i>	<i>(ba:ṇ)^(x)</i>
Abujhmaria	VI	<i>b:or</i>	<i>bad</i>		<i>b:orlo:r</i>	<i>baw</i>		= Set I <i>bo:r</i>	= Set I <i>bad,</i> <i>ba:ta:</i>
Gondi		(a) <i>ayen, eyen</i> (b) <i>inikan</i>	(a) <i>ayed</i> (b) <i>inika(d)</i>		(a) <i>ayer</i> (b) <i>iniker</i>	(a) <i>ayek</i> (b) <i>inikeṇ</i>		?	?
Konda ^(xi)	VI	(a) <i>estanju</i> (b) <i>ananju, inanju</i>	(a) <i>estari</i> (NOM) (a) <i>estaṛin-</i> (other cases)	(a) <i>estari</i> (NOM) (a) <i>estaṛ-</i> (other cases)	(a) <i>estaru</i> (b) <i>anaru, inaru</i>	(a) <i>estai</i> (NOM) (a) <i>estaskan-</i> (other cases)			<i>imbai, embai</i> , = Set I <i>ana(r)i,</i> <i>ba:r^(xiii)</i> <i>ina(r)i^(xii)</i>
Kui ^(xi)	VI								

Language	GS ⁽ⁱ⁾	Set I						Set II	
		SG			PL			HUM (‘who?’)	NON<HUM> (‘what?’)
		M	F	N	M	F	N		
		<div>(b) <i>anari</i>, <i>inari</i> (NOM) no oblique all cases^{xii)} cases</div> <div>(a) <i>emminasi</i> (a) <i>emminayi</i> (b) <i>amba(:)si</i>, <i>imba^ʔasi</i> (b) <i>amba(:)yi</i>, <i>imbai</i> (a) <i>emminari</i> (a) <i>emmina^ʔi</i> (b) <i>amba(:)ri</i>, (b) <i>amba(:)^ʔi</i>, <i>imbai</i>, <i>imbaari</i> <i>imba^ʔari</i>, <i>e:nari</i> <i>amnar</i> (a) <i>innakar</i> (a) <i>innakadek</i> (a) <i>innakaj</i> (b) <i>inakar</i>, (b) <i>indek</i> <i>iner</i>^{xvi)} (b) ? (not recorded)</div>							
Kuvi ^(xi)	VI ^(xiv)								
Manda	(X)								
Pengo	IV								

Language	GS ⁽¹⁾	Set I						Set II	
		SG			PL			HUM (‘who?’)	NON<HUM> (‘what?’)
		M	F	N	M	F	N		
Naiki	VI	<i>e:n</i>	<i>e:d</i>		<i>e:r</i>		<i>e:v</i>	= Set I <i>e:n</i>	<i>ta:(n)</i>
Gadaba	VI	<i>eyiŋ, e:yŋd</i>	<i>e:di</i>		<i>e:yir, e:yr</i>		<i>e:vi</i>	= Set I <i>e:yr</i>	<i>e:na(co:y)</i>
Parji	VI	(a) <i>e:d</i> (NOM), other cases: <i>e:r-</i> ^{xviii}	—		(a) <i>e:r</i> ^{xviii}		—	= Set I <i>e:d</i> (NOM) other cases: <i>e:r-</i> ^{xviii} , (predica- tive rarely as: <i>na:t-ed</i> (SG), <i>na:t-er</i> (PL))	<i>na:(to)</i>
		(b) (predicative (usual): <i>na:t-ot</i> & rare: <i>na:t-ed</i>)			(b) (predicative & rare: <i>na:t-er</i>)		(b) (predicative & rare: <i>na:t-ov</i>)		
Northern Dravidian									
Malpaharia	(IX)	<i>ike-h</i> (NOM)	<i>ike-θ</i> (NOM)			<i>ike-r</i>		<i>ne:d(u)</i> ^{xix}	<i>indr(du)</i>
Malto		no oblique cases	all cases: <i>ike-</i>		no oblique cases		(all cases: <i>ike-</i>)		
Sauria Malto ^{xi}	VI	(a) <i>ike-h</i> (NOM)	(a) <i>iki-θ</i> (NOM)	(a) <i>iku-θ</i> (NOM)		(a) <i>ikike-r(-)</i>		?	<i>indru</i> (cf. Set I)
		(b) <i>ne:re-h</i> (NOM)	(b) <i>ne:(ri)-θ</i> (NOM)	(b) predicative: <i>indr-θ</i> (NOM)	(b) <i>ne:re-r</i> (NOM), oblique cases: <i>ne:me:re-r-</i>		(b) <i>indrindru-r-</i>		
		oblique cases: <i>nek-</i>							
Kurux	V/IV	—	—	—	—	—	—	<i>ne:</i> ^{xx}	<i>endr(a:),</i> <i>e:kda:</i> ^{xx}
Brahui	VIII	<i>ara:(d)</i>				<i>ara:fk</i>		<i>de(r)</i>	<i>ant</i>

NB: The interrogatives are presented in the nominative. In the other cases the same distinctions are made, unless explicitly mentioned otherwise. The original orthography of the sources has been generally preserved intact. A dot below, as in *ṇ*, indicates that the consonant is retroflex, while a macron below, as in *n̄*, usually indicates that the consonant is alveolar. It seems that usually the alveolar pronunciation is explicitly marked only when a corresponding dental is available in a given language.

Notes:

- i) The type of the gender system according to Table III.3:11.
- ii) In Modern Tamil the interrogative *evār* functions honorifically, i.e. it is not used when the speaker believes that “the person in question is a (low-status) male or female” (Shiffman 1999). This interrogative does not have a corresponding plural form and when used as a subject it requires the 3SG.HON[NON<N>] marker *-ār* to be used on the verb (Pillai & Kothandaraman 1972:350-351). It seems that when a honorific meaning needs to be expressed in the plural, the general interrogative *ya:r* ‘who?’ can be used instead.
- iii) The verb in a clause with *ya:r* as the subject is usually said to take either the 3SG.HON[NON<N>] marker *-ār* or the 3PL.NON<N> marker *-ārkaḷ* (Rangan, Suseela & Rajendran 2002), although the 3SG.HON[NON<N>] marker seems to be the default choice. However, as far as I can judge from an example in Annamalai & Steever (1998:117), the 3SG.M marker *-āṇ* is also possible.
- iv) The verb in a clause with the interrogative *ya:ru* as the subject takes 3PL.NON<N> marking (cf. Bhat 1989:372).
- v) In old texts also *e:n/enna(tu)* (Andronov 1996:90-91).
- vi) Interestingly, although Tulu lacks separate M.SG and F.SG interrogative pronominals, it makes this distinction in demonstrative-cum-anaphoric pronominals and nouns. The interrogatives *da:nē* and *ja:nē* are mentioned by Andronov (1978:276) and Burrow & Emeneau (1984:466) but not in Bhat’s (1998) description of Tulu.
- vii) According to Ramanarasimham (1998:187-188), in Old Telugu the separate F.SG and N.SG interrogatives *evvate* and *e:mi* function only as predicate nominals and are then obligatory. This seems to imply that the NON<M>.SG interrogative *e:di* is used pronominally and non-predicatively both in questions about humans (viz. women) and non-humans. Furthermore, *e:(mi)*, but not *evvate*, can be used attributively as ‘which [N]’, what [N]?’ irrespective of the number/gender of the N (Ramanarasimham 1998:195). Interestingly, it seems that in Modern Telugu *e:m(i)* has evolved into a Set II interrogative ‘what?’, which can only be used non-predicatively, as in (1), while predicative and attributive functions are executed by the interrogative *e:miti/enti/etti* ‘what (kind of)?’, as in (2).

Modern Telugu (South-Central Dravidian; India; <http://www.languageshome.com/English-Telugu.htm>)

- | | | | | | | |
|-----|--------------|----------------------|---------------------|-----|------------------------|--|
| (1) | <i>ne:nu</i> | <i>e:mi</i> | <i>ceyy-a:li?</i> | (2) | <i>praṇṇalu</i> | <i>enti?</i> |
| | 1SG.NOM | IPW[S2].NON<HUM>.NOM | do-should | | question.PL.NON<M>.NOM | IPW[S2].NON<HUM>.NOM |
| | | | ‘What should I do?’ | | | ‘What are/were/will be the question(s)?’ |

viii) In Modern Telugu the interrogative *ewaru* also functions as honorific (cf. note ⁱⁱ above) (Krishnamurti & Gwynn 1985:72) and when used as a subject, it co-occurs with NON<N>.PL marking on the verb (Krishnamurti 1998:229). It is not clear whether the interrogatives *ewate* F.SG and *e:di* NON<M>.SG are then explicitly non-honorific when used in questions about humans.

ix) The interrogative *bor* also functions as honorific (cf. note ⁱⁱ above), but as far as I can judge from Steever (1998a), it still requires M.PL marking on the verb.

x) The interrogatives *ba:ṇ* and *ba:ta-* are mentioned by Burrow & Emeneau (1984:466) but not in Steever’s (1998a) description of Muria Gondi. Furthermore, according to Natarajan (1985:330), *ba:ṇ* is another variant of the M.SG form in Muria Gondi.

xi) The forms under (a) are the so-called “definite” interrogative pronominals, usually glossed as ‘which man (etc.)?’, and the forms under (b) are the so-called “indefinite” interrogative pronominals, usually glossed as ‘what man (etc.)?’ and sometimes as ‘what kind of man (etc.)?’ (see Section III.3.3.2).

xii) According to Winfield (1928:46), in the genitive, accusative and ablative the number distinction of the interrogative pronominal *anari/inari*, *anai/inai* tends to be neutralized in favour of the plural form *ana/ina*. The distinction is preserved in the nominative and dative.

xiii) According to Winfield (1928:45), when used as a subject, *imbai/embai/ba:i* is usually followed by a verb with a NON<M>.SG marker (e.g., *-ne* in the present and *-te* in the past).

xiv) Interestingly, although Kuvi lacks separate M.PL and F.PL interrogative pronominals and verbal markers, it does make this distinction in the demonstrative-cum-anaphoric pronouns (cf. Reddy 1979:66).

xv) There may exist a plural form *e:nai/imbinaɪ* as well (Burrow & Emeneau 1984:467), but it is not mentioned in Reddy’s (1979) description of Kuvi.

xvi) Note that in Pengo the opposition between the SG.M and the PL.M forms is optional and can be made only in the nominative. Elsewhere, the plural form has been generalized.

xvii) “Kolami (a)” forms come from Subrahmanyam (1998:308) and “Kolami (b)” forms from Emeneau (1955:55, 164-171). Strangely enough, both sources describe the same dialect, Wardha Kolami, and Subrahmanyam’s description is said to be “based on Emeneau’s (1955) treatment of the

Wardha dialect” (1998:301), although apparently expanded with the author’s own unpublished data. Note that Emeneau (1955:55) distinguishes special feminine forms both in the singular and the plural (although the singular form “does not occur in [his] material”) and that he does not describe the forms beginning in *ta(:)n-* as exclusively predicative but as interrogatives “substituting for non-persons only (? things only, rather than animals)”. The latter feature brings the interrogative system of Emeneau’s variety of Kolami closer to Naiki.

^{xviii}) Note that in Parji the opposition between the SG.M and the PL.M forms has been preserved only in the nominative. Elsewhere, the plural form has been generalized.

^{xix}) Steever (1998b:368) labels *ne:d(u)* as “human plural” when presenting case paradigms of the interrogative pronominals. However, on the same page he claims that “*ne:du* is always singular” and in an example given by Steever *ne:d(u)* is used with a M.SG marker *-(a)h* on the verb.

^{xx}) According to (Grierson 1906:414; Hahn 1911:29-30), although *ne:* is used only to inquire about humans, when used as a subject it is followed by a verb with a NON<M>.SG marker, *-i:* in the present and *-a:* in the past, as in *ne: barch-a:* ‘Who came?’. According to (Burrow & Emeneau 1984:467), *ne:* and *endr* may be reduplicated to express the idea of plurality. However, Hahn (1911:30-31) gives this possibility only for *ne:* and only as a means to form a free choice indefinite *ne:ne:* ‘whosoever’. The latter is followed by a verb with a M.PL marker, as in *ne:ne: barch-ar* ‘Whosoever came’. The three non-human interrogatives are distinguished by Hahn (1911:30-33) as follows. On the whole, the interrogatives *endr* and *endra:*, both mean ‘what?’, but the former tends to be used in a more “abstract sense”, that is rather about actions and situations (3a), while the latter is used “only definitely”, that is rather about entities, thoughts, words, etc. (3b). The interrogative *endr* can also be used attributively as ‘what (kind of)?’. The non-human interrogative *e:kda:* is claimed to be used only as ‘which one (non-human)?’.

Kurux (Northern Dravidian; India. Hahn 1911:31)

- | | | | | | | | |
|-----|----|----------------------|-----------------------|--|----|----------------------|-----------------------|
| (3) | a. | <i>endr</i> | <i>ra’i:?</i> | | b. | <i>endra:</i> | <i>ra’i:?</i> |
| | | IPW[S2].NON<HUM>.NOM | remain.PRS.3SG.NON<M> | | | IPW[S2].NON<HUM>.NOM | remain.PRS.3SG.NON<M> |
| | | ‘What’s the matter?’ | | | | ‘What is it?’ | |

References

- [no author]. 1927. *Language hand-book: Savara*. Calcutta: Tea Districts Labour Association.
http://ling.lll.hawaii.edu/austroasiatic/AA/savara_handbook.pdf.
- Abu-Absi, Samir. 1995. *Chadian Arabic*. München: LINCOM Europa.
- Achab, Karim. 2005. Le système de genre et son origine en berbère et en chamito-sémitique. In Antoine Lonnet & Amina Mettouchi (eds.), *Les langues chamito-sémitiques (afro-asiatiques)*, Vol. 1, 97-128. Paris: Ophrys.
- Adelaar, Willem F. H. & (with the collaboration of) Pieter C. Muysken. 2004. *The languages of the Andes*. Cambridge: Cambridge University Press.
- Aikhenvald, Alexandra Y. 1984. *Структурно-типологическая классификация берберских языков [A structural and typological classification of Berber languages]*. Moscow: Institute of Oriental Studies PhD thesis.
- Aikhenvald, Alexandra Y. 1994. Classe nominal e gênero nas línguas Aruák. *Boletim do Museu Paraense Emílio Goeldi* 10(2). 137-259.
- Aikhenvald, Alexandra Y. 1995a. *Bare*. München: LINCOM Europa.
- Aikhenvald, Alexandra Y. 1995b. Split ergativity in Berber languages. *St. Petersburg Journal of African Studies* 4. 39-68.
- Aikhenvald, Alexandra Y. 1998. Warekena. In Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*, Vol. 4, 225-440. Berlin: Mouton de Gruyter.
- Aikhenvald, Alexandra Y. 1999. The Arawak language family. In R.M.W. Dixon & Alexandra Y. Aikhenvald (eds.), *The Amazonian languages*, 65-106. Cambridge: Cambridge University Press.
- Aikhenvald, Alexandra Y. 2000. *Classifiers: A typology of noun categorization devices*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra Y. & Alexander Y. Militarev. 1991. Ливийско-гуанчские языки [Lybia-Guanche languages]. In Igor M. Diakonoff & G. Sh. Sharbatov (eds.), *Языки Азии и Африки [Languages of Asia and Africa]*, Vol. 4.2: Афразийские языки [Afro-Asiatic languages], 148-267. Moscow: Nauka.
- Aikhenvald, Alexandra Y. & R.M.W. Dixon. 1999. Other small families and isolates. In R.M.W. Dixon & Alexandra Y. Aikhenvald (eds.), *The Amazonian languages*, 341-383. Cambridge: Cambridge University Press.
- Akhmedov, German. 1997. Classification of interrogative sentences in Lezgian language. In Mikhail Ye. Alekseev (ed.), *Дагестанский лингвистический сборник: Памяти Г.А. Климова [Dagestan papers in linguistics: In memory of G.A. Klimov]*, Vol. 4, 114-119. Moscow: Academia.
- Alekseev, Mikhail Ye. 1985. *Вопросы сравнительно-исторической грамматики лезгинских языков. Морфология. Синтаксис [Issues in the comparative-historical grammar of the Lezgian languages. Morphology. Syntax]*. Moscow: Nauka.

- Alekseev, Mikhail Ye. 1988. *Сравнительно-историческая морфология аваро-андийских языков* [*The comparative-historical morphology of the Avar-Andic languages*]. Moscow: Nauka.
- Alekseev, Mikhail Ye. 2001a. Цезские языки [Tsezic languages]. In Mikhail Ye. Alekseev, Georgy A. Klimov, Sergey A. Starostin & Yakov G. Testelelets (eds.), *Языки мира: Кавказские языки* [*Languages of the World: Caucasian languages*], 299-303. Moscow: Academia.
- Alekseev, Mikhail Ye. 2001b. Рутульский язык [Rutul]. In Mikhail Ye. Alekseev, Georgy A. Klimov, Sergey A. Starostin & Yakov G. Testelelets (eds.), *Языки мира: Кавказские языки* [*Languages of the World: Caucasian languages*], 409-420. Moscow: Academia.
- Alekseev, Mikhail Ye. 2001c. Нахско-дагестанские языки [Nakh-Dagestanian languages]. In Viktor V. Vinogradov et al. (ed.), *Языки Российской Федерации и соседних государств: Энциклопедия* [*The languages of Russia and adjacent states: An encyclopedia*], Vol. 2: K-R, 329-330. Moscow: Nauka.
- Alekseev, Mikhail Ye. 2001d. Нахско-дагестанские языки [Nakh-Dagestanian languages]. In Mikhail Ye. Alekseev, Georgy A. Klimov, Sergey A. Starostin & Yakov G. Testelelets (eds.), *Языки мира: Кавказские языки* [*Languages of the World: Caucasian languages*], 156-166. Moscow: Academia.
- Alekseev, Mikhail Ye. 2001e. Лезгинские языки [Lezgian languages]. In Mikhail Ye. Alekseev, Georgy A. Klimov, Sergey A. Starostin & Yakov G. Testelelets (eds.), *Языки мира: Кавказские языки* [*Languages of the World: Caucasian languages*], 369-373. Moscow: Academia.
- Allan, Robin, Philip Holmes & Tom Lundskaer-Nielsen. 2003. *Danish: a comprehensive grammar*. London: Routledge.
- Allen, J. H., (edited by) James Bradstreet Greenough, G. L. Kittredge, A. A. Howard & Benj. L. D'Ooge. 1903. *New Latin grammar for schools and colleges founded on comparative grammar*. Boston: Ginn.
- Allen, James Paul. 2000. *Middle Egyptian: An introduction to the language and culture of hieroglyphs*. Cambridge: Cambridge University Press.
- Ambrazas, Vytautas (ed.). 1997. *Lithuanian grammar*. Vilnius: Baltos lankos.
- Anderbeck, Karl. 2006. An initial reconstruction of Proto-Lampungic: Phonology and basic vocabulary. *Tenth International Conference on Austronesian Linguistics, 17-20 January 2006*, Puerto Princessa City, Palawan, Philippines.
<http://www.sil.org/asia/philippines/ical/papers.html>.
- Anderson, Gregory D.S. 2001. A new classification of South Munda: evidence from comparative verb morphology. *Indian Linguistics* 62. 21-36.
- Anderson, Ronald Jaime. 1985-1986. *Cuentos folklóricos de los ashéninka*. Yarinacocha: Instituto Lingüístico de Verano.
- Andronov, Mikhail S. 1978. *Сравнительная грамматика дравидийских языков* [*A comparative grammar of the Dravidian languages*]. Moscow: Nauka.
- Andronov, Mikhail S. 1980. *The Brahui language*. Moscow: Nauka.

- Andronov, Mikhail S. 1996. *A grammar of Malayalam in historical treatment*. Wiesbaden: Harrassowitz.
- Angenot-de Lima, Geralda. 2002. *Description phonologique, grammaticale et lexicale du moré, langue amazonienne de Bolivie et du Brésil*. Leiden: Universiteit Leiden PhD Thesis.
- Annamalai, E. & Sanford B. Steever. 1998. Modern Tamil. In Sanford B. Steever (ed.), *The Dravidian languages*, 100-128. London: Routledge.
- Antunes de Araujo, Gabriel. 2004. *A grammar of Sabanê, a Nambikwaran language*. Amsterdam: Vrije Universiteit Amsterdam PhD thesis.
- Applegate, Joseph R. 1958. *An outline of the structure of Shilha*. New York: American Council of Learned Societies.
- Appleyard, D.L. 1987. A grammatical sketch of Khamtanga - I. *Bulletin of the School of Oriental and African Studies* 50(2). 241-266.
- Araújo, Gabriel. 2000. *Fonologia e morfologia da língua Maxakali*. Campinas: Universidade Estadual de Campinas MA thesis.
- Asher, R.E. & T.C. Kumari. 1997. *Malayalam*. London: Routledge.
- Austin, Peter. 1980 (ms.). *Some notes on the grammar of Djiwarli*.
- Avrorin, Valentin A. 2000. *Грамматика маньчжурского письменного языка [A grammar of Written Manchu]*. St.Petersburg: Nauka.
- Aze, F. Richard & Trish Aze. 1973. Parengi texts. In Ronald L. Trail (ed.), *Patterns in clause, sentence, and discourse in selected languages of India and Nepal*, Vol. 3: Texts, 213-362. Kathmandu: Summer Institute of Linguistics & Institute of Nepal and Asiatic Studies, Tribhuvan University.
- Babakaev, Vasily D. 1980. *Очерки морфологической структуры ассамского языка [Studies in the morphological structure of the Assamese language]*. Moscow: Nauka.
- Badawi, Elsaid, Michael G. Carter & Adrian Gully. 2004. *Modern Written Arabic: A comprehensive grammar*. London: Routledge.
- Bailly, Anatole. 1901. *Abrégé du dictionnaire grec-français*. Paris: Hachette.
<http://home.tiscali.be/tabularium/bailly/index.html>.
- Baker, Brett J. 1999. *Word structure in Ngalakan*. Sidney: University of Sidney PhD thesis.
- Bakker, Dik. to appear. Language sampling. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology*. Oxford: Oxford University Press.
- Bala, G.G., (edited by) Bonny Sands & Will Grundy. 1998. *Hadza stories and songs (translated by Bonny Sands)*. Los Angeles: Friends of the Hadzabe.
- Bamgboṣe, Ayọ. 1980. Questions and roles: The example of Yoruba. *Afrika und Übersee* LXIII. 21-38.
- Baptista, Priscilla & Ruth Wallin. 1967. Baure. In Esther Matteson (ed.), *Bolivian Indian grammars*, Vol. 1, 27-84. Norman: Summer Institute of Linguistics of the University of Oklahoma.
- Bárcena, Alonso & (edited by) Samuel A. Lafone Quevedo M.A. 1893. *Arte de la lengua Toba*. La Plata: Talleres de Publicaciones del Museo de La Plata.

- Barlaan, Rodolfo R. 1977. Fact and fiction in Isneg narrative discourse. *Studies in Philippine Linguistics* 1(2). 110-142.
- Barth, Jacob. 1913. *Die Pronominalbildung in den semitischen Sprachen*. Leipzig: Hinrichs.
- Barulin, A.N. 1980. Некоторые проблемы семантического анализа вопросительных местоимений (на материале русского языка) [Some problems of semantic analysis of interrogative pronouns (on the example of Russian)]. In I.F. Vardul' (ed.), *Теория и типология местоимений [Theoretical and typological studies of pronouns]*, 27-49. Moscow: Nauka.
- Basset, André. 1952. *La langue berbère*. London: Oxford University Press.
- Basset, René. 1909. *Mission au Senegal*. Vol. 1: Étude sur le dialecte zénaga. Notes sur le hassaniya. Recherches historiques sur les maures. Paris: Ernest Leroux.
- Bastin, Yvonne, André Coupez & Michael Mann. 1999. *Continuity and divergence in the Bantu languages: perspectives from a lexicostatistic study*. Tervuren: Koninklijk Museum voor Midden-Afrika.
- Beard, Thomas. 2003. Syntax. In Derek Nurse & Gérard Philippson (eds.), *The Bantu Languages*, 121-142. London & New York: Routledge.
- Bell, Alan. 1978. Language samples. In Joseph H. Greenberg (ed.), *Universals of human language*, Vol. 1: Method and theory, 123-156. Stanford: Stanford University Press.
- Ben Mamou, El Arbi. 2005a. *Dialecte berbère de Tamezret (Tunisie): Lexique berbère-français*. http://www.atmazret.info/atmazret_info/ATMAZRAT_DOCS/ugmazret.pdf.
- Ben Mamou, El Arbi. 2005b. *Dialecte berbère de Tamezret (Tunisie): Exemples berbère-français*. http://www.atmazret.info/atmazret_info/ATMAZRAT_DOCS/eddwi.pdf.
- Bender, Lionel M. 1990. A survey of Omotic grammemes. In Philip Baldi (ed.), *Linguistic change and reconstruction methodology*, 661-696. Berlin: Mouton de Gruyter.
- Bender, Lionel M. 2000. *Comparative morphology of the Omotic languages*. München: LINCOM Europa.
- Benjamin, Geoffrey. 1976a. Austroasiatic subgroupings and prehistory in the Malay Peninsula. In Philip N. Jenner, Laurence C. Thompson & Stanley Starosta (eds.), *Austroasiatic studies*, Vol. 1, 37-128. Honolulu: The University Press of Hawaii.
- Benjamin, Geoffrey. 1976b. An outline of Temiar grammar. In Philip N. Jenner, Laurence C. Thompson & Stanley Starosta (eds.), *Austroasiatic studies*, Vol. 1, 129-188. Honolulu: The University Press of Hawaii.
- Bentolila, Fernand. 1981. *Grammaire fonctionnelle d'un parler berbère: Aït Seghrouchen d'Oum Jeniba (Maroc)*. Paris: SELAF.
- Besnier, Niko. 2000. *Tuvaluan*. London: Routledge.
- Bhaskararao, Peri. 1980. *Konekor Gadaba: A Dravidian language*. Pune: Deccan College.
- Bhaskararao, Peri & S. K. Joshi. 1985. A study of Newari classifiers. *Bulletin of the Deccan College Research Institute* 44. 17-31.
- Bhat, Darbhe Narayana Shankara. 1989. *An introduction to Indian grammars*. Vol. 2: Wh-words. Mysore: Central Institute of Indian Languages.

- Bhat, Darbhe Narayana Shankara. 1998. Tulu. In Sanford B. Steever (ed.), *The Dravidian languages*, 158-180. London: Routledge.
- Bhat, Darbhe Narayana Shankara. 2000. The indefinite-interrogative puzzle. *Linguistic Typology* 4(3). 365-400.
- Bhat, Darbhe Narayana Shankara. 2004. *Pronouns: A cross-linguistic study*. Oxford: Oxford University Press.
- Biarnay, Samuel. 1908. *Étude sur le dialecte berbère de Ouargla*. Paris: Ernest Leroux.
- Bittner, Maximilian. 1916. *Studien zur Šhauri-Sprache in den Bergen von Dofâr am Persischen Meerbusen*. Vol. 2: Zum Verbum und zu den ubrigen Redeteilen. Wien: Hölder.
- Bittner, Maximilian. 1917. *Studien zur Šhauri-Sprache in den Bergen von Dofâr am Persischen Meerbusen*. Vol. 3: Zu ausgewählten Texten. Wien: Hölder.
- Black, Mary B. 1971. On Ojibwa question constructions. *International Journal of American Linguistics* 37(3). 146-151.
- Black, Paul. 1976. Werizoid. In Lionel M. Bender (ed.), *The Non-Semitic languages of Ethiopia*, 222-231. East Lansing: Michigan State University.
- Blake, Barry J. 2003. *The Bunganditj (Buwandik) language of the Mount Gambier Region*. Canberra: Pacific Linguistics.
- Blake, Frank R. 1906. Contributions to comparative Philippine grammar. *Journal of the American Oriental Society* 27. 317-396.
- Bliese, Loren. 1976. Afar. In Lionel M. Bender (ed.), *The Non-Semitic languages of Ethiopia*, 133-165. East Lansing: Michigan State University.
- Bloch, Oscar & Walther von Wartburg. 1968. *Dictionnaire étymologique de la langue française*. 5th edn. Paris: Presses Universitaires de France.
- Blust, Robert, Russell Gray & Simon Greenhill. 2006. *Austronesian Basic Vocabulary Database*. <http://language.psy.auckland.ac.nz/austronesian/>.
- Bokarev, Yevgeny A. 1959. Цезские (дидойские) языки Дагестана [*The Tsezic (Didoic) languages of Dagestan*]. Moscow: The Academy of Sciences of the USSR.
- Bokarev, Yevgeny A. 1967a. Хваршинский язык [Khvarshi]. In Viktor V. Vinogradov *et al.* (ed.), *Языки народов СССР [Languages of the peoples of the USSR]*, Vol. 4: Иберийско-кавказские языки [Ibero-Caucasian languages], 421-435. Moscow: Nauka.
- Bokarev, Yevgeny A. 1967b. Цезский язык [Tsez]. In Viktor V. Vinogradov *et al.* (ed.), *Языки народов СССР [Languages of the peoples of the USSR]*, Vol. 4: Иберийско-кавказские языки [Ibero-Caucasian languages], 404-420. Moscow: Nauka.
- Bokarev, Yevgeny A. 1967c. Гинухский язык [Hinukh]. In Viktor V. Vinogradov *et al.* (ed.), *Языки народов СССР [Languages of the peoples of the USSR]*, Vol. 4: Иберийско-кавказские языки [Ibero-Caucasian languages], 436-454. Moscow: Nauka.
- Bokarev, Yevgeny A. 1967d. Гунзибский язык [Hunzib]. In Viktor V. Vinogradov *et al.* (ed.), *Языки народов СССР [Languages of the peoples of the USSR]*, Vol. 4: Иберийско-кавказские языки [Ibero-Caucasian languages], 472-487. Moscow: Nauka.
- du Bourguet, Pierre. 1976. *Grammaire fonctionnelle et progressive de l'égyptien démotique*. Leuven: Peeters.

- Bradshaw, Joel. 2001. Iwal grammar essentials, with comparative notes. In Andrew Pawley, Malcolm Ross & Darrell Tryon (eds.), *The boy from Bundaberg: Studies in Melanesian linguistics in honour of Tom Dutton*, 51–74. Canberra: Pacific Linguistics.
- Bradshaw, Joel & Francisc Czobor (eds.). 2004. *Otto Dempwolff's grammar of the Jabêm language in New Guinea*. Honolulu: University of Hawai'i Press.
- von Brandenstein, C.G. 1980. *Ngadjumaja: An Aboriginal language of south-east Western Australia*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck. [Document 0552 of the Aboriginal Studies Electronic Data Archive, Australian Institute of Aboriginal and Torres Strait Islander Studies, Canberra, typed by Catriona Hyslop, 1994].
- Breen, Gavan. 1974. *Wakaya grammar*. [Document 0047 of the Aboriginal Studies Electronic Data Archive, Australian Institute of Aboriginal and Torres Strait Islander Studies, Canberra, typed by Jane Simpson, 1985].
- Breen, Gavan. 2003. Wanyi and Garrwa comparative data. In Nicholas Evans (ed.), *The non-Pama-Nyungan languages of northern Australia: comparative studies of the continent's most linguistically complex region*, 425–462. Canberra: Pacific Linguistics.
- Brockelmann, Carl. 1908. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*. Vol. 1: Laut- und Formenlehre. Berlin: Reuther und Reichard.
- Brockelmann, Carl. 1913. *Grundriss der vergleichenden Grammatik der semitischen Sprachen*. Vol. 2: Syntax. Berlin: Reuther und Reichard.
- Bruening, Benjamin. 2001. *Syntax at the edge: Cross-clausal phenomena and the syntax of Passamaquoddy*. Cambridge: Massachusetts Institute of Technology PhD thesis.
- Buccellati, Giorgio. 1997. Akkadian. In Robert Hetzron (ed.), *The Semitic languages*, 69–99. London: Routledge.
- Bulatova, Nadezhda Ya. & Lenor Grenoble. 1999. *Evenki*. München: LINCOM Europa.
- Burling, Robbins. 1961. *A Garo grammar*. Poona: Deccan College.
- Burling, Robbins. 2004. *Language of the Modhupur Mandi*. Vol. 2: The lexicon. New Delhi: Bibliophile South Asia. <http://www-personal.umich.edu/~rburling/Vol-II.pdf>.
- Burling, Robbins. forthcoming. *Comparative phonology of the Boro-Garo languages*. Mysore: Central Institute of Indian Languages. <http://www-personal.umich.edu/~rburling/ProtoB-GCogChart.xls>.
- Burridge, Kate. 1996. Yulparija sketch grammar. In William B. McGregor (ed.), *Studies in Kimberley languages in honour of Howard Coate*, 15–70. München: LINCOM Europa.
- Burrow, Thomas & Sudhibhushan Bhattacharya. 1953. *The Parji language: A Dravidian language of Bastar*. Hertford: Stephen Austin and Sons.
- Burrow, Thomas & Sudhibhushan Bhattacharya. 1970. *The Pengo language: Grammar, texts, and vocabulary*. Oxford: Clarendon Press.
- Burrow, Thomas & Murray B. Emeneau. 1984. *A Dravidian etymological dictionary*. 2nd edn. Oxford: Oxford University Press.
- Butcher, Andrew & Nicholas Reid. 1989. *Redressing the balance on Australian stop contrasts: comments on Austin's (1988) "Phonological voicing contrasts in Australian Aboriginal languages"*. <http://www.flinders.edu.au/speechpath/AustinCrit.pdf>.

- Bybee, Joan L., Revere D. Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: Chicago University Press.
- de Calassanti-Motylinski, Gustave Adolphe. 1904. *Le dialecte berbère de R'edames*. Paris: Ernest Leroux.
- de Calassanti-Motylinski, Gustave Adolphe & René Basset. 1908. *Grammaire, dialogues et dictionnaire touaregs*. Vol. 1: Grammaire et dictionnaire français-touareg. Algiers: Pierre Fontana.
- Callender, John B. 1975. *Middle Egyptian*. Malibu: Undena.
- Campbell, Lyle & Terrence Kaufman. 1985. Mayan linguistics: Where are we now? *Annual Review of Anthropology* 14. 187-198.
- Capell, Arthur. 1942. Languages of Arnhem Land. *Oceania* 12(4). 364-392.
- Capell, Arthur. 1943. Languages of Arnhem Land. *Oceania* 13(1). 24-50.
- Capell, Arthur. 1953. Notes on the Waramunga language, Central Australia. *Oceania* 23(4). 297-311.
- Capell, Arthur. 1984. The Laragia language, *Papers in Australian linguistics*, Vol. 16, 55-106. Canberra: Pacific Linguistics.
- Cauquelin, Josiane. 1991. The Puyuma language. *Bijdragen tot de Taal-, Land- en Volkenkunde* 147(1). 17-60.
- Černý, Jaroslav. 1976. *Coptic etymological dictionary*. Cambridge: Cambridge University Press.
- Černý, Jaroslav. 1978. *A late Egyptian grammar*. Rome: Biblical Institute Press.
- Chadwick, Neil. 1975. *A descriptive grammar of the Djingili language*. Canberra: Australian Institute of Aboriginal Studies.
- Chapman, Shirley. 1986. Paumarí interrogatives. In Joseph E. Grimes (ed.), *Sentence initial devices*, 215-233. Arlington: Summer Institute of Linguistics.
- Chapman, Shirley. 1988. Interrogativos paumarí. *Série Lingüística* 9(2). 89-116.
- Chapman, Shirley & Desmond C. Derbyshire. 1991. Paumarí. In Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*, Vol. 3, 161-352. Berlin: Mouton de Gruyter.
- Chisholm, Williams S., Louis T. Milic & John A.C. Greppin (eds.). 1984. *Interrogativity: A colloquium on the grammar, typology and pragmatics of questions in seven diverse languages*. Amsterdam: John Benjamins.
- Chkadua, Lidiya P. & R.N. Klychev. 2001. Абазинский язык [Abaza]. In Mikhail Ye. Alekseev, Georgy A. Klimov, Sergey A. Starostin & Yakov G. Testelet (eds.), *Языки мира: Кавказские языки [Languages of the World: Caucasian languages]*, 131-146. Moscow: Academia.
- Choi, Jane. 2005. A corpus-based approach to the Korean wh-question words *mwusun* and *enu*. In Robert Bley-Vroman & HyunSook Ko (eds.), *Corpus linguistics for Korean language learning and teaching*, 107-147. Honolulu: National Foreign Language Resource Center, University of Hawai'i.

- Cincius, Vera I. *et al.* (ed.). 1975-1977. *Сравнительный словарь тунгусо-маньчжурских языков: материалы к этимологическому словарю* [*A comparative dictionary of the Tungus-Manchu languages: Materials toward an etymological dictionary*]. Vol. 1-2. Leningrad: Nauka.
- Clendon, Mark. 1994. *A learner's grammar of Worora*. Broome: Kimberley Language Resource Centre.
- Coate, H.H.J. & Lynette F. Oates. 1970. *A grammar of Ngarinjin (Western Australia)*. Canberra: Australian Institute of Aboriginal Studies.
- Cohen, David, Marie-Claude Simeone-Senelle & Martine Vanhove. 2002. The grammaticalization of 'say' and 'do': An areal phenomenon in East Africa. In Tom Güldemann & Manfred von Roncador (eds.), *Reported Discourse: A meeting ground for different linguistic domains*, 227-251. Amsterdam: John Benjamins.
- Conti Rossini, Carlo. 1913. *Schizzo del dialetto Saho dell'Alta Assaorta in Eritrea*. Rome: Reale accademia dei Lincei.
- Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2005. Sex-based and non-sex-based gender systems. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, 130-134. Oxford: Oxford University Press.
- Cowan, Marion M. 1956. *Aprendamos castellano. Una gramática castellana-tzotzil*. Mexico: Summer Institute of Linguistics.
- Cowan, Marion M. & William R. Merrifield. 1968. The verb phrase in Huixtec Tzotzil. *Language* 44(2). 284-305.
- Croft, William. 2003. *Typology and universals*. 2nd edn. Cambridge: Cambridge University Press.
- Crum, Walter Ewing. 1962 [1939]. *A Coptic dictionary*. Oxford: Clarendon.
- Cyr, Danielle. 1992. Montagnais. In Jacques Muraire (ed.), *Les langues autochtones du Québec*, 247-286. Québec: Les publications du Québec.
- Cysouw, Michael. 2003. Against implicational universals. *Linguistic Typology* 7(1). 89-101.
- Cysouw, Michael. 2004a. Minimal systems of interrogative words. *Workshop on Questions*, MPI-EVA, Leipzig. <http://email.eva.mpg.de/~cysouw/pdf/cysouwHANDMINQ.pdf>.
- Cysouw, Michael. 2004b. Interrogative words: an exercise in lexical typology. *Bantu grammar: description and theory*, 3, ZAS Berlin. <http://email.eva.mpg.de/~cysouw/pdf/cysouwQUESTION.pdf>.
- Cysouw, Michael. 2005. Quantitative methods in typology. In Reinhard Köhler, Gabriel Altmann & Rajmund G. Piotrowski (eds.), *Quantitative linguistics: An international handbook*, 554-577. Berlin: Mouton de Gruyter.
- Cysouw, Michael. forthcoming. Content interrogatives in Asheninca Campa (Arawak): corpus study and typological comparison. *International Journal of American Linguistics*.

- Dahl, Östen & Kari Fraurud. 1996. Animacy in grammar and discourse. In Thorstein Fretheim & Jeanette K. Gundel (eds.), *Reference and referent accessibility*, 47-64. Amsterdam: John Benjamins.
- Dahl, Otto Christian. 1976 [1973]. *Proto-Austronesian*. London: Curzon Press.
- Dambriūnas, Leonardas, Antanas Klimas & William R. Schmalstieg. 1972. *Introduction to Modern Lithuanian*. Brooklyn: Franciscan Fathers.
- Darmesteter, Arsène & (edited by) Léopold Sudre. 1902. *Cours de grammaire historique de la langue française*. 4th edn. Paris: Delagrave.
- Das, A. Sisir Kumar. 1973. *Structure of Malto*. Annamalainagar: Annamalai University.
- Dasgupta, Dipankar. 1978. *Linguistic studies in Juang, Kharia Thar, Lodha Mal-Pahariya, Ghatoali Pahariya*. Calcutta: Anthropological Survey of India, Government of India.
- Dayley, Jon P. 1985. *Tzutujil grammar*. Berkeley: University of California Press.
- Dayley, Jon P. 1989. *Timbisha (Panamint) Shoshone grammar*. Berkeley: University of California Press.
- Delgaty, Colin. 1964. *Vocabulario tzotzil de San Andrés, Chiapas*. Mexico: Summer Institute of Linguistics.
- Dempwolff, Otto. 1934-1938. *Vergleichende Lautlehre des austronesischen Wortschatzes*. (3 Vols.). Berlin: Dietrich Reimer.
- Dempwolff, Otto. 1939. *Grammatik der Jabêm-Sprache auf Neuguinea*. Hamburg: Friederichsen de Gruyter & Co.
- Dench, Alan. 1991. Panyjima. In R.M.W. Dixon & Barry J. Blake (eds.), *The handbook of Australian languages*, Vol. 4, 125-243. Oxford: Oxford University Press.
- Dench, Alan. 1998. *Yingkarta*. München: LINCOM Europa.
- Dentan, Robert K. & (edited by) Thomas Doyle. 2003 [1963]. *Preliminary field notes on the Semai language*. <http://www.keene.edu/library/OrangAsli/DentanBaer.pdf>.
- Depuydt, Leo. 1999. *Fundamentals of Egyptian grammar*. Vol. 1: Elements. Norton: Frog.
- Diakonoff, Igor M. 1965. *Семитохамитские языки: Опыт классификации [The Semito-Hamitic languages: An essay in classification]*. Moscow: Nauka.
- Diakonoff, Igor M. 1988. *Afrasian languages*. Moscow: Nauka.
- Dienhart, John M. 1997. *The Mayan languages: a comparative vocabulary (electronic version)*. Odense: Odense University. <http://maya.hum.sdu.dk>.
- Diessel, Holger. 1999. *Demonstratives: form, function, and grammaticalization*. Amsterdam: John Benjamins.
- Diessel, Holger. 2003. The relationship between demonstratives and interrogatives. *Studies in Language* 27(3). 635-655.
- Diffloth, Gérard. 2005. The contribution of linguistic palaeontology to the homeland of Austro-Asiatic. In Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds.), *The peopling of East Asia: putting together archaeology, linguistics and genetics*, 77-81. London: Routledge Curzon.
- Dik, Simon. 1989. *The theory of functional grammar*. Vol. 1: The structure of the clause. Dordrecht: Foris publications.

- Dixon, R.M.W. 1977. *A grammar of Yidin*. Cambridge: Cambridge University Press.
- Dixon, R.M.W. 1980. *The languages of Australia*. Cambridge: Cambridge University Press.
- Dixon, R.M.W. 1997. *The rise and fall of languages*. Cambridge: Cambridge University Press.
- Dixon, R.M.W. 2002. *Australian languages: their nature and development*. Cambridge: Cambridge University Press.
- Doerfer, Gerhard. 1978. Classification problems of Tungus. In Gerhard Doerfer & Michael Weiers (eds.), *Tungusica*, Vol. 1, 1-26. Wiesbaden: Otto Harrassowitz.
- Dolgopolsky, Aron B. 1991. Кушитские языки [Cushitic languages]. In Igor M. Diakonoff & G. Sh. Sharbatov (eds.), *Языки Азии и Африки [Languages of Asia and Africa]*, Vol. 4.2: Афризийские языки [Afro-Asiatic languages], 3-147. Moscow: Nauka.
- Donaldson, Tamsin. 1980. *Ngiyambaa, the language of the Wangaaybuwan*. Cambridge: Cambridge University Press.
- Donegan, Patricia J. & David Stampe. 2004. *Munda Lexical Archive*.
<http://www.ling.hawaii.edu/austroasiatic/AA/Munda/Dictionaries/munda-archive.zip>.
- Doneux, Jean L. 1971. Deux études sur les interrogatifs en zone J du bantou. *Africana linguistica* 5. 119-138.
- Doneux, Jean L. & Claire Grégoire. 1977. Une séquence vocalique à la finale de quelques thèmes en protobantou. *Africana linguistica* 7. 159-199.
- Douglas, Wilfrid H. 1981. Watjarri. In R.M.W. Dixon & Barry J. Blake (eds.), *The handbook of Australian languages*, Vol. 2, 197-273. Amsterdam: John Benjamins.
- Dryer, Matthew S. 1989a. Plural words. *Linguistics* 27. 865-895.
- Dryer, Matthew S. 1989b. Large linguistic areas and language sampling. *Studies in Language* 13(2). 257-292.
- Dryer, Matthew S. 2000. Counting genera vs. counting languages. *Linguistic Typology* 4(3). 334-350.
- Dryer, Matthew S. 2003. Significant and non-significant implicational universals. *Linguistic Typology* 7(1). 108-128.
- Du Feu, Veronica. 1996. *Rapanui*. London: Routledge.
- Dumestre, Gérard. 2003. *Grammaire fondamentale du bambara*. Paris: Karthala.
- Dumézil, Georges. 1933. *Introduction à la grammaire comparée des langues caucasiennes du nord*. Paris: Librairie Ancienne Honoré Champion.
- Dymshic, Zalman M. 1962. *Язык урду [The Urdu language]*. Moscow: Nauka.
- Eastlack, Charles L. 1968. Terena (Arawakan) pronouns. *International Journal of American Linguistics* 34(1). 1-8.
- Eberle, Andrea & (in collaboration with) Regine Schulz. 2004. *Koptisch: ein Leitfaden durch das Saïdische*. München: LINCOM Europa.
- Ebert, Karen H. 1996. *Koḍava*. München: LINCOM Europa.
- Eckert, Paul & Joyce Hudson. 1988. *Wangka Wiru: A handbook for the Pitjantjatjara language learner*. Adelaide: South Australian College of Advanced Education.
- Edel, Elmar. 1955. *Altägyptische Grammatik*. Vol. 1. Rome: Pontificium institutum biblicum.
- Edel, Elmar. 1964. *Altägyptische Grammatik*. Vol. 2. Rome: Pontificium institutum biblicum.

- Edmondson, Jerold A. 1995. English-Bolyu glossary. *Mon-Khmer Studies* 23. 133-159.
- Ehrenreich, Paul M.A. 1896. Ein Beitrag zur Charakteristik der botokudischen Sprache, *Festschrift für Adolf Bastian zu seinem 70. Geburtstage gewidmet*, 607-630. Berlin: Otto Elsner.
- Ehret, Christopher. 1995. *Reconstructing Proto-Afroasiatic (Proto-Afrasian): vowels, tone, consonants, and vocabulary*. Berkeley: University of California Press.
- Elanskaya, Alla I. 1991. Коптский язык [The Coptic language]. In Igor M. Diakonoff & G. Sh. Sharbatov (eds.), *Языки Азии и Африки [Languages of Asia and Africa]*, Vol. 4.2: *Афразийские языки [Afro-Asiatic languages]*, 298-322. Moscow: Nauka.
- Emeneau, Murray B. 1955. *Kolami: A Dravidian language*. Berkeley: University of California Press.
- Emeneau, Murray B. 1971. *Toda songs*. Oxford: Clarendon.
- Emeneau, Murray B. 1984. *Toda grammar and texts*. Philadelphia: American Philosophical Society.
- Endzelin, J. 1923. *Lettische Grammatik*. Heidelberg: Carl Winter.
- Erman, Adolf. 1968 [1933]. *Neuägyptische Grammatik*. Hildesheim: Olms.
- Evans, Nicholas. 1994. Kayardild. In Cliff Goddard & Anna Wierzbicka (eds.), *Semantic and lexical universals: theory and empirical findings*, 203-228. Amsterdam: John Benjamins.
- Evans, Nicholas. 1995. *A grammar of Kayardild: with historical-comparative notes on Tangkic*. Berlin: Mouton de Gruyter.
- Evans, Nicholas. 1997. Head classes and agreement classes in the Mayali dialect chain. In Mark Harvey & Nicholas Reid (eds.), *Nominal classification in Aboriginal Australia*, 105-146. Amsterdam: John Benjamins.
- Evans, Nicholas. 2003a. Comparative non-Pama-Nyungan and Australian historical linguistics. In Nicholas Evans (ed.), *The non-Pama-Nyungan languages of northern Australia: comparative studies of the continent's most linguistically complex region*, 3-25. Canberra: Pacific Linguistics.
- Evans, Nicholas. 2003b. *Bininj Gun-wok: a pan-dialectal grammar of Mayali, Kunwinjku and Kune*. Canberra: Pacific Linguistics.
- Evans, Nicholas (ed.). 2003c. *The non-Pama-Nyungan languages of northern Australia: comparative studies of the continent's most linguistically complex region*. Canberra: Pacific Linguistics.
- Evans, Nicholas. 2005. Australian languages reconsidered: a review of Dixon (2002). *Oceanic Linguistics* 44(1). 242-286.
- Evans, Nicholas & Rhys Jones. 1997. The cradle of the Pama-Nyungans: archaeological and linguistic speculations. In Patrick McConvell & Nicholas Evans (eds.), *Archaeology and linguistics: Aboriginal Australia in global perspective*, 385-417. Melbourne: Oxford University Press.
- Everett, Daniel L. & Barbara Kern. 1997. *Wari': the Pacaas Novos language of western Brazil*. London: Routledge.

- van den Eynde, Karel & Piet Mertens. 2003. La valence: l'approche pronominale et son application au lexique verbal. *Journal of French Language Studies* 13(1). 63-104.
- Faber, Alice. 1991. The diachronic relationship between negative and interrogative markers in Semitic. In Alan S. Kaye (ed.), *Semitic studies in honor of Wolf Leslau: On the occasion of his eighty-fifth birthday, November 14th, 1991*, 411-429. Wiesbaden: Otto Harrassowitz.
- Faidherbe, Louis Léon César. 1877. *Le zénaga des tribus sénégalaises*. Paris: Ernest Leroux.
- Favreau, R., J. Péret, M. Valière, B. Horiot, J. Biard-Millérioux, J. Pitié & B. Guesnier. 1983. *Poitou*. Le Puy-en-Velay: Christine Bonneton, Éditeur.
- Feldpausch, Tom & Becky Feldpausch. 1992. *Namia grammar essentials*. Ukarumpa via Lae: Summer Institute of Linguistics.
- Ferguson, Judith, Cari Hollinger, Linda Criswell & Nancy L. Morse. 2000. El cubeo. In María Stella González de Pérez & María Luisa Rodríguez de Montes (eds.), *Lenguas indígenas de Colombia: una visión descriptiva*, 357-372. Bogotá: Instituto Caro y Cuervo.
- Ferlus, Michel. 1974. Délimitation des groupes linguistiques austroasiatiques dans le Centre-Indochinois. *Asie du Sud-Est et Monde Insulindien* 5(1). 15-23.
- Ferreirinho, Naomi. 1993. *Selected topics in the grammar of Limos Kalinga, the Philippines*. Canberra: Pacific Linguistics.
- Ferrell, Raleigh. 1969. *Taiwan aboriginal groups: Problems in cultural and linguistic classification*. Nankang, Taipei: Institute of Ethnology, Academia Sinica.
- Fischer, Wolfdietrich. 1997. Classical Arabic. In Robert Hetzron (ed.), *The Semitic languages*, 187-219. London: Routledge.
- Fleming, Ilah. 1966. Carib. In Marvin K. Mayers (ed.), *Languages of Guatemala*, 303-308. The Hague: Mouton & Co.
- Fortier, A. 1972 [1895]. *Louisiana folktales in French dialect and English translation*. New York: Kraus Reprints.
- de Foucauld, Charles. 1951-1952. *Dictionnaire touareg-français: dialecte de l'Ahaggar*. Impr. nationale de France.
- Frajzyngier, Zygmunt. 1985. Interrogative sentences in Chadic. *Journal of West African Languages* 15(1). 57-72.
- Francis, David A., Sr. & Robert M. Leavitt. 2007. *Maliseet-Passamaquoddy dictionary*. Fredericton: Mi'kmaq-Maliseet Institute at the University of New Brunswick.
<http://www.lib.unb.ca/Texts/Maliseet/dictionary>.
- François, Alexandre. forthcoming. Noun articles in Torres and Banks languages: Conservation and innovation. In Jeff Siegel, John Lynch & Diana Eades (eds.), *Linguistic description and linguistic applications: Studies in memory of Terry Crowley*. New York: John Benjamins.
- Frank, Paul Stephen. 1985. *A grammar of Ika*. Philadelphia: University of Pennsylvania Press.
- Frantz, Donald G. 1991. *Blackfoot grammar*. Toronto: University of Toronto Press.
- Furby, Christine E. 1972. The pronominal system of Garawa. *Oceanic Linguistics* 11(1). 1-31.
- Galand, Lionel. 1974. Présentation d'un parler: le touareg de l'Aïr (Introduction grammaticale). In Petites Sœurs de Jésus (ed.), *Contes touaregs de l'Aïr*, 13-41. Paris: SELAF.

- Galand, Lionel. 2002 [1974]. Présentation d'un parler: le touareg de l'Air (Introduction grammaticale). In Lionel Galand (ed.), *Études de linguistique berbère*, 117-146. Leuven: Peeters.
- Galucio, Ana Vilacy. 2001. *The morphosyntax of Mekens (Tupi)*. Chicago: University of Chicago PhD thesis.
- García Matzar, Pedro & José Obispo Rodríguez Guaján. 1997. *Rukemik ri Kaqchikel chi' / Gramática Kaqchikel*. Guatemala: Cholsamaj.
- Gatschet, Albert S. 1877. The Timucua language. *Proceedings of the American Philosophical Society* 16. 626-642.
- Gatschet, Albert S. 1878. The Timucua language. *Proceedings of the American Philosophical Society* 17. 490-504.
- Gatschet, Albert S. 1880. The Timucua language. *Proceedings of the American Philosophical Society* 18. 465-502.
- Genko, A.N. 1955. *Абазинский язык: Грамматический очерк наречия Тапанта [Abaza: A grammatical description of the Tapanta dialect]*. Moscow: The Academy of Sciences of the USSR.
- Gerzenstein, Ana. 1994. *Lengua maká: estudio descriptivo*. Buenos Aires: Universidad de Buenos Aires.
- Gesenius, Wilhelm. 1846. *Gesenius' Hebrew and Chaldee lexicon to the Old Testament Scriptures (translated by Samuel Prideaux Tregelles)*. <http://www.blueletterbible.org>.
- Gesenius, Wilhelm & E. Kautzsch. 1962. *Wilhelm Gesenius' hebräische Grammatik (völlig umgearbeitet von E. Kautzsch)*. 28th edn. Hildesheim: Georg Olms.
- GICLI = Grupo de Investigação Científica de Línguas Indígenas. *Swadesh lists of Brazilian native languages*. <http://paginas.terra.com.br/educacao/GICLI>.
- Gieser, C. Richard. 1963. *A grammatical sketch of Kalinga*. Nasuli, Malaybalay, Bukidnon: Summer Institute of Linguistics.
- Giridhar, P.P. 1980. *Angami grammar*. Mysore: Central Institute of Indian Languages.
- Givón, Talmy. 1984. Ute. In Williams S. Chisholm, Louis T. Milic & John A.C. Greppin (eds.), *Interrogativity: A Colloquium on the Grammar, Typology and Pragmatics of Questions in Seven Diverse Languages*. Amsterdam: John Benjamins.
- Glasgow, Kathleen. 1984. Burarra word classes, *Papers in Australian linguistics*, Vol. 16, 1-54. Canberra: Pacific Linguistics.
- Goddard, Cliff. 1983. *A semantically-oriented grammar of the Yankunytjatjara dialect of the Western Desert language*. Canberra: Australian National University.
- Goddard, Cliff. 1994. Lexical primitives in Yankunytjatjara. In Cliff Goddard & Anna Wierzbicka (eds.), *Semantic and lexical universals: theory and empirical findings*, 229-262. Amsterdam: John Benjamins.
- Goddard, Cliff. 2001. Lexico-semantic universals: A critical overview. *Linguistic Typology* 5(1). 1-65.
- de Goeje, C.H. 1928. *The Arawak language of Guiana*. Amsterdam: Koninklijke Akademie van Wetenschappen.

- Gomes dos Santos, Manoel. 2006. *Uma gramática do Wapixana (Aruák) - aspectos da fonologia, da morfologia e da sintaxe*. Campinas: Universidade Estadual de Campinas PhD thesis.
- Gordon, Cyrus H. 1997. Amorite and Eblaite. In Robert Hetzron (ed.), *The Semitic languages*, 100-113. London: Routledge.
- Gordon, Raymond G., Jr. (ed.). 2005. *Ethnologue: Languages of the world*. 15th edn. Dallas: SIL International. Online version: <http://www.ethnologue.com/>.
- Gorgoniyev, Y.A. 1966. *The Khmer language*. Moscow: Nauka.
- Gragg, Gene. 1976. Oromo of Wellegga. In Lionel M. Bender (ed.), *The Non-Semitic languages of Ethiopia*, 166-195. East Lansing: Michigan State University.
- Granberry, Julian. 1990. A grammatical sketch of Timucua. *International Journal of American Linguistics* 56(1). 60-101.
- Granberry, Julian. 1993. *A grammar and dictionary of the Timucua language*. 3rd edn. Tuscaloosa: University of Alabama Press.
- Green, Rebecca. 1995. *A grammar of Gurr-goni (North Central Arnhem Land)*. Canberra: Australian National University PhD thesis.
- Greenberg, Joseph H. 1966a. *Language universals*. The Hague: Mouton & Co.
- Greenberg, Joseph H. 1966b. *The languages of Africa*. Bloomington: Indiana University.
- Greenberg, Joseph H. 1978. How does a language acquire gender markers? In Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik (eds.), *Universals of human language*, Vol. 3: Word structure, 47-82. Stanford: Stanford University Press.
- Grevisse, Maurice & (edited by) André Goosse. 1986. *Le bon usage: grammaire française*. 12th edn. Paris: Duculot.
- Grierson, George Abraham (ed.). 1906a. *Linguistic survey of India*. Vol. 1, Part 2: Comparative vocabulary of Indian languages. Calcutta: Superintendent of Government Printing.
- Grierson, George Abraham (ed.). 1906b. *Linguistic survey of India*. Vol. 2: Mōn-Khmēr and Siamese-Chinese families (including Khassi & Tai). Calcutta: Superintendent of Government Printing.
- Grierson, George Abraham (ed.). 1906c. *Linguistic survey of India*. Vol. 4: Muṇḍā and Dravidian languages. Calcutta: Superintendent of Government Printing.
- Griffiths, Glyn. 1976. O elemento interrogativo em Kadiwéu. *Série Lingüística* 6. 35-41.
- Grondona, Verónica María. 1998. *A grammar of Mocoví*. Pittsburgh: University of Pittsburgh PhD thesis.
- Gruber, Jeffrey S. 1975. Bushman languages of the Kalahari: =Hòǎ - Vocabulary - Stems, =Hòǎ - Vocabulary - Recorded utterances. Washington, D.C.: Technical Project Report to the National Endowment for the Humanities.
- Gudava, T.E. 1967. Ботлихский язык [Botlikh]. In Viktor V. Vinogradov *et al.* (ed.), *Языки народов СССР [Languages of the peoples of the USSR]*, Vol. 4: Иберийско-кавказские языки [Ibero-Caucasian languages], 293-306. Moscow: Nauka.
- Guilbeau, J. 1950. *The French spoken in La Fourche Parish, Louisiana*. Raleigh: University of North Carolina PhD thesis.

- Guillaume, Antoine. 2004. *A grammar of Cavineña, an Amazonian language of northern Bolivia*. Melbourne: La Trobe University PhD thesis.
- Güldemann, Tom. 2005. *Studies in Tuu (Southern Khoisan)*. Leipzig: Institut für Afrikanistik, Universität Leipzig.
- Güldemann, Tom. 2006. Noun categorization, agreement, and “janus-headed” compounds in Western !Xõo. *Rara & Rarissima - Collecting and interpreting unusual characteristics of human languages*, 29 March - 1 April, Max Planck Institute for Evolutionary Anthropology, Leipzig.
http://email.eva.mpg.de/~gueldema/pdf/GenusHU_Berlin05Handout.pdf.
- Güldemann, Tom. forthcoming. Greenberg’s “case” for Khoisan: the morphological evidence. In Rainer Voßen & Dymitr Ibriszimow (eds.), *Problems of linguistic-historical reconstruction in Africa*. Köln: Rüdiger Köppe.
- Güldemann, Tom & Rainer Vossen. 2000. Khoisan. In Bernd Heine & Derek Nurse (eds.), *African languages: an introduction*, 99-122. Cambridge: Cambridge University Press.
- Gutt, Ernst-August. 1997. The Silte group (East Gurage). In Robert Hetzron (ed.), *The Semitic languages*, 509-534. London: Routledge.
- Häberl, Charles. 2006. *The Neo-Mandaic dialect of Khorramshahr*: Harvard University PhD thesis.
- Haeseryn, W., K. Romijn, G. Geerts, J. de Rooij & M.C. van den Toorn. 1997. *Algemene Nederlandse Spraakkunst*. 2nd edn. Groningen/Deurne: Martinus Nijhoff uitgevers/Wolters Plantyn.
- Hahn, Ferdinand. 1911. *Grammar of the Kurukh language*. Calcutta: Bengal Secretariat Press. [Reprinted in no date, Delhi: Mittal Publications].
- Hamouma, Hamid. 1987. *Manuel de grammaire berbère (kabyle)*. Paris: Association de Culture Berbère.
- Hanawalt, Charlie. 2006. Bitter or sweet? The vital role of sociolinguistic survey in Lampungic dialectology. *Tenth International Conference on Austronesian Linguistics, 17-20 January 2006*, Puerto Princesa City, Palawan, Philippines.
<http://www.sil.org/asia/philippines/ical/papers.html>.
- Hanawalt, Charlie, Kristina Tarp & The Liana Husain. in preparation (ms.). *The Lampungic groups of South Sumatra and Lampung: A dialectology and sociolinguistic study*.
- Hanke, Wanda. 1960. Die Sprache der Kaišana am Lago Mapari. *Zeitschrift für Ethnologie* 85. 54-81.
- Hanoteau, Adolphe. 1896. *Essai de grammaire de la langue tamacheck’*. 2nd edn. Algiers: Adolphe Jourdan.
- Hansen, Kenneth C. & Lesley E. Hansen. 1978. *The core of Pintupi grammar*. Alice Springs: Institute for Aboriginal Development, Summer Institute of Linguistics.
- Hargreaves, David. 2003. Kathmandu Newar (Nepāl Bhāṣā). In Graham Thurgood & Randy J. LaPolla (eds.), *The Sino-Tibetan languages*, 371-384. London: Routledge.

- Harvey, Mark. 1997a. Nominal classification and gender in Aboriginal Australia. In Mark Harvey & Nicholas Reid (eds.), *Nominal classification in Aboriginal Australia*, 17-62. Amsterdam: John Benjamins.
- Harvey, Mark. 1997b. Head and agreement classes: an areal perspective. In Mark Harvey & Nicholas Reid (eds.), *Nominal classification in Aboriginal Australia*, 147-164. Amsterdam: John Benjamins.
- Harvey, Mark. 2001. *A grammar of Limilngan: a language of the Mary River Region, Northern Territory, Australia*. Canberra: Pacific Linguistics.
- Harvey, Mark. 2002. *A grammar of Gaagudju*. Berlin: Mouton de Gruyter.
- Harvey, Mark. 2003. Reconstruction of pronominals among the non-Pama-Nyungan languages. In Nicholas Evans (ed.), *The non-Pama-Nyungan languages of northern Australia: comparative studies of the continent's most linguistically complex region*, 475-513. Canberra: Pacific Linguistics.
- Harvey, Mark & Nicholas Reid (eds.). 1997. *Nominal classification in Aboriginal Australia*. Amsterdam: John Benjamins.
- Harvey, Mark, Ian Green & Rachel Nordlinger. 2006. From prefixes to suffixes: Typological change in Northern Australia. *Diachronica* 23(2). 289-311.
- Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford: Clarendon.
- Haspelmath, Martin, Matthew S. Dryer, David Gil & Bernard Comrie (eds.). 2005. *The world atlas of language structures*. Oxford: Oxford University Press.
- Hayashi, Makoto & Kyung-eun Yoon. 2006. A cross-linguistic exploration of demonstratives in interaction: With particular reference to the context of word-formulation trouble. *Studies in Language* 30(3). 485-540.
- Heath, Jeffrey. 1978. *Ngandi: grammar, texts and dictionary*. Canberra: Australian Institute of Aboriginal Studies.
- Heath, Jeffrey. 1981. *Basic materials in Mara: grammar, texts and dictionary*. Canberra: Pacific Linguistics.
- Heath, Jeffrey. 1984. *Functional grammar of Nunggubuyu*. Canberra: Australian Institute of Aboriginal Studies.
- Heath, Jeffrey. 2005. *A grammar of Tamashek (Tuareg of Mali)*. Berlin: Mouton de Gruyter.
- Hercus, Luise Anna. 1999. *A grammar of the Wirangu language from the west coast of South Australia*. Canberra: Pacific Linguistics.
- Hettick, Donna & Carolyn Kent. 1967. Uses of the particle *di* [in Northern Kankanay]: Library of the Summer Institute of Linguistics, Manila.
- Hettick, Donna & Judy Wallace. 1978. Northern Kankanay texts. *Studies in Philippine Linguistics* 2(2). 1-28.
- Hetzron, Robert. 1997. *The Semitic languages*. London: Routledge.
- Himmelmann, Nikolaus P. 1996. Demonstratives in narrative discourse: A taxonomy of universal uses. In A.B. Fox (ed.), *Studies in anaphora*, 205-254. Amsterdam: John Benjamins.

- Hjelmslev, Louis. 1956 [1971]. Animé et inanimé, personnel et non-personnel. In Louis Hjelmslev (ed.), *Essais linguistiques*, 211-249. Paris: Les Éditions de Minuit.
- Hoch, James Erich. 1994. *Semitic words in Egyptian texts of the New Kingdom and the Third Intermediate period*. Princeton: Princeton University Press.
- Hoddinott, William G. & Frances M. Kofod. 1988. *The Ngankikurungkurr language (Daly River area, Northern Territory)*. Canberra: Pacific Linguistics.
- Hohulin, Dick & Austin Hale. 1977. Physical, temporal and cognitive location in Keley-i demonstratives. *Studies in Philippine Linguistics* 1(2). 209-230.
- Holst, Jan Henrik. 2001. *Lettische Grammatik*. Hamburg: Helmut Buske.
- Hornborg, Alf. 2005. Ethnogenesis, regional integration, and ecology in prehistoric Amazonia: Toward a system perspective. *Current Anthropology* 46(4). 589-620.
- Hosokawa, Komei. 1991. *The Yawuru language of West Kimberley: a meaning-based description*. Canberra: Australian National University PhD thesis.
- Houtzagers, Peter (ed.). 2003. *Russische grammatica*. Bussum: Uitgeverij Coutinho.
- Huang, Lillian M. 1996. Interrogative constructions in Mayrinax Atayal. *Bulletin of National Taiwan Normal University* 41. 263-296.
- Huang, Lillian M. 2006. Manifestations of participants in Atayal: A cross-dialectal study. *Tenth International Conference on Austronesian Linguistics, 17-20 January 2006*, Puerto Princessa City, Palawan, Philippines. <http://www.sil.org/asia/philippines/ical/papers.html>.
- Huang, Lillian M., Elizabeth S. Zeitoun, Marie M. Yeh, Anna H. Chang & Joy J. Wu. 1999. Interrogative constructions in some Formosan languages. In Yuen-mei Yin, I-li Yang & Hui-chen Chan (eds.), *Chinese languages and linguistics*, Vol. V: Interactions in language, 639-680. Taipei: Institute of Linguistics, Preparatory Office, Academia Sinica.
- Huang, Shuping & Li-May Sung. 2005. The undergoer focus *Ma-* in Kavalan. *Taiwan-Japan Joint Workshop on Austronesian Languages, June 23-24, 2005*, Graduate Institute of Linguistics, National Taiwan University, Linguistic Society of Taiwan. <http://homepage.ntu.edu.tw/~gilntu/data/workshop%20on%20Austronesian/6%20huang%20&%20sung.pdf>.
- Hudson, Grover. 1976a. Highland East Cushitic. In Lionel M. Bender (ed.), *The Non-Semitic languages of Ethiopia*, 232-277. East Lansing: Michigan State University.
- Hudson, Grover. 1997. Amharic and Argobba. In Robert Hetzron (ed.), *The Semitic languages*, 457-485. London: Routledge.
- Hudson, Richard A. 1976b. Beja. In Lionel M. Bender (ed.), *The Non-Semitic languages of Ethiopia*, 97-132. East Lansing: Michigan State University.
- Huff, Wayne & Alice Huff. 1971. Uspantec. In Mary Shaw (ed.), *According to our ancestors: Folk texts from Guatemala and Honduras*, 241-246, 501-507. Norman: Summer Institute of Linguistics of the University of Oklahoma.
- Huffman, Franklin. 1971 (ms.). *Mon-Khmer vocabulary lists*. <http://sealang.net/archives/huffman/>.
- Hulstaert, Gustaaf. 1938. *Praktische grammatica van het Lonkundo (Lomongo) (Belgisch Kongo)*. Antwerpen: De Sikkel.

- Hulstaert, Gustaaf. 1957. *Dictionnaire Lomongo-Français*. Tervuren: Musée royal du Congo belge.
- Hulstaert, Gustaaf. 1961. *Grammaire du lomongo*. Vol. 1: La phonologie. Tervuren: Musée royal d'Afrique centrale.
- Hulstaert, Gustaaf. 1965. *Grammaire du lomongo*. Vol. 2: La morphologie. Tervuren: Musée royal d'Afrique centrale.
- Hulstaert, Gustaaf. 1966. *Grammaire du lomongo*. Vol. 3: La syntaxe. Tervuren: Musée royal d'Afrique centrale.
- Hulstaert, Gustaaf. 1987. Bonkándá wă Njakomba (Elaká ěa Yôóko)/ Le Nouveau Testament en lomongo. Mbandaka: Imprimatur Mbandaka.
- Hulstaert, Gustaaf. 1993. Les interrogatifs dans les dialectes Móngo. *Annales Aequatoria* 14. 345-377.
- Idiatov, Dmitry. 2006. Rare and nonexistent interrogative pro-word types: interrogative pro-verbs & Co. *Rara & Rarissima*, 29 March - 1 April 2006, Max Planck Institute for Evolutionary Anthropology, Leipzig.
- Ingham, Bruce. 1973. Urban and rural Arabic in Khūzistān. *Bulletin of the School of Oriental and African Studies* 36(3). 533-553.
- Ingham, Bruce. 1982. *North East Arabian dialects*. London & Boston: Kegan Paul International.
- Izre'el, Shlomo. 1998. *Canaanite-Akkadian*. München: LINCOM Europa.
- Izre'el, Shlomo & Eran Cohen. 2004. *Literary Old Babylonian*. München: LINCOM Europa.
- Jacq, Pascale & Paul Sidwell. 1999. *Sapuan (Səpuar)*. München: LINCOM Europa.
- Janda, Laura A. 1996. Figure, ground and animacy in Slavic declension. *The Slavic and East European Journal* 40(2). 325-355.
- Janssen, Dirk P., Balthasar Bickel & Fernando Zúñiga. 2006. Randomization tests in language typology. *Linguistic Typology* 10(3). 419-440.
- Janzen, Hermann. 1976. Structure and function of clauses and phrases in Pale. In Philip N. Jenner, Laurence C. Thompson & Stanley Starosta (eds.), *Austroasiatic studies*, Vol. 1, 669-691. Honolulu: The University Press of Hawaii.
- Jastrow, Otto. 1997. The Neo-Aramaic languages. In Robert Hetzron (ed.), *The Semitic languages*, 334-377. London: Routledge.
- Jespersen, Otto. 1927. *A modern English grammar on historical principles*. Heidelberg: Carl Winter.
- Johnson, Janet H. 2000. *Thus wrote 'Onchsheshonqy - An introductory grammar of Demotic*. 3rd edn. Chicago: Oriental Institute, University of Chicago.
<http://oi.uchicago.edu/OI/DEPT/PUB/SRC/SAOC/45/SAOC45.html>.
- Johnson, Janet H. (ed.). 2001. *The Demotic dictionary of the Oriental Institute of the University of Chicago*. Chicago: Oriental Institute, University of Chicago.
<http://oi.uchicago.edu/OI/DEPT/PUB/SRC/CDD/CDD.html>.
- Johnson, Mark. 1987. *The body in the mind*. Chicago: University of Chicago Press.
- Johnstone, T. M. 1975. *The modern South Arabian languages*. Malibu: Undena.

- Joüon, Paul. 1965 [1923]. *Grammaire de l'hébreu biblique*. Rome: Institut biblique pontifical.
- Junge, Friedrich. 1996. *Einführung in die Grammatik des Neuägyptischen*. Wiesbaden: Harrassowitz.
- Karlberg, Göran. 1954. *The English interrogative pronouns: a study of their syntactic history*. Stockholm: Almqvist & Wiksell.
- Kaufman, Stephen A. 1997. Aramaic. In Robert Hetzron (ed.), *The Semitic languages*, 114-130. London: Routledge.
- Kaufman, Terrence. 1976. *Proyecto de alfabetos y ortografías para escribir las lenguas mayances*. Guatemala: Proyecto lingüístico Francisco Marroquín.
- Kaufman, Terrence & (with the assistance of) John Justeson. 2003. *A preliminary Mayan etymological dictionary*. <http://www.famsi.org/reports/01051/pmed.pdf>.
- Kaye, Alan S. & Judith Rosenhouse. 1997. Arabic dialects and Maltese. In Robert Hetzron (ed.), *The Semitic languages*, 263-311. London: Routledge.
- Keen, Sandra. 1983. Yukulta. In R.M.W. Dixon & Barry J. Blake (eds.), *Handbook of Australian languages*, Vol. 3, 191-305. Amsterdam: John Benjamins.
- Khalilov, Madzhid Sh. 2001. Цезский язык [Tsez]. In Mikhail Ye. Alekseev, Georgy A. Klimov, Sergey A. Starostin & Yakov G. Testelelets (eds.), *Языки мира: Кавказские языки* [*Languages of the World: Caucasian languages*], 320-331. Moscow: Academia.
- Khalilov, Madzhid Sh. 2005. Цезский язык [Tsez]. In Viktor V. Vinogradov et al. (ed.), *Языки Российской Федерации и соседних государств: Энциклопедия* [*The languages of Russia and adjacent states: An encyclopedia*], Vol. 3: S-Ya, 346-354. Moscow: Academia.
- Kite, Suzanne & Stephen A. Wurm. 2004. *The Duunjdjawa language of southeast Queensland: grammar, texts and vocabulary*. Canberra: Pacific Linguistics.
- Klein, Harriet E. Manelis. 2001. *Toba*. München: LINCOM Europa.
- Klokeid, Terry J. 1976. *Topics in Lardil grammar*. Cambridge: Massachusetts Institute of Technology PhD thesis.
- Klychev, R.N. & Lidiya P. Chkadua. 2001. Абхазский язык [Abkhaz]. In Mikhail Ye. Alekseev, Georgy A. Klimov, Sergey A. Starostin & Yakov G. Testelelets (eds.), *Языки мира: Кавказские языки* [*Languages of the World: Caucasian languages*], 113-131. Moscow: Academia.
- Konstantinova, O.A. 1968. Эвенкийский язык [Evenki]. In Viktor V. Vinogradov et al. (ed.), *Языки народов СССР* [*Languages of the peoples of the USSR*], Vol. 5: Монгольские, тунгусо-маньчжурские и палеоазиатские языки [Mongolic, Tungus-Manchu and Paleoasiatic languages], 68-87. Leningrad: Nauka.
- Korostovtsev, Mikhail A. 1973. *Grammaire du néo-égyptien*. Moscow: Nauka.
- Korostovtsev, Mikhail A. 1991. Среднеегипетский язык [The Middle Egyptian language]. In Igor M. Diakonoff & G. Sh. Sharbatov (eds.), *Языки Азии и Африки* [*Languages of Asia and Africa*], Vol. 4.2: Афразийские языки [Afro-Asiatic languages], 268-297. Moscow: Nauka.
- Kossmann, Maarten G. 1997. *Grammaire du parler berbère de Figuig (Maroc oriental)*. Paris: Peeters.

- Krishnamurti, Bhadriraju. 1998. Telugu. In Sanford B. Steever (ed.), *The Dravidian languages*, 202-240. London: Routledge.
- Krishnamurti, Bhadriraju & J.P.L. Gwynn. 1985. *A grammar of Modern Telugu*. Delhi: Oxford University Press.
- Kruspe, Nicole D. 1999. *Semelai*. Melbourne: The University of Melbourne PhD thesis.
- Kühner, Raphael & Friedrich Holzweissig. 1974 [1912]. *Ausführliche Grammatik der lateinischen Sprache*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Kühner, Raphael, (edited by) Friedrich Blass & Bernhard Gerth. 1890-1904. *Ausführliche Grammatik der griechischen Sprache*. Hannover: Hahn.
- Lambdin, Thomas O. 1983. *Introduction to Sahidic Coptic*. Macon: Mercer University Press.
- Landaburu, Jon. 2000. La lengua ika. In María Stella González de Pérez & María Luisa Rodríguez de Montes (eds.), *Lenguas indígenas de Colombia: una visión descriptiva*, 733-748. Bogotá: Instituto Caro y Cuervo.
- Landin, David J. 1984. An outline of the syntactic structure of Karitiâna sentences. In Robert A. Dooley (ed.), *Estudos sobre línguas tupi do Brasil*, 219-254. Brasília: Summer Institute of Linguistics.
- Landin, David J. 2005. *Dicionário e léxico Karitiana/Português*. 2nd edn. Cuiabá: Sociedade Internacional de Lingüística.
<http://www.sil.org/americas/brasil/PUBLICNS/DICTGRAM/KTDic.pdf>.
- Landin, Rachel M. 1987. Conjunções karitiâna de nível superior. *Série Lingüística* 9(1). 199-206.
- Lane, G. 1935. Notes on Louisiana French II: Spoken Standard French of Saint Martinville. *Language* 10. 323-333.
- Landry, Jacques (in collaboration with Jean-Marie Dallet). 1968. *Ghadamès: Étude linguistique et ethnographique*. Vol. 1: Textes, notes philologiques et ethnographiques. Fort-National (Algeria): Fichier de Documentation Berbère.
- Landry, Jacques. 1973. *Ghadamès: Étude linguistique et ethnographique*. Vol. 2: Glossaire (parler des Ayt Waziten). Algiers: Le Fichier Périodique.
- Lapenda, Geraldo. 1968. *Estrutura da língua iatê (falada pelos índios Fulniôs em Pernambuco)*. Recife: Universidade Federal de Pernambuco.
- Lavency, Marius. 1985. *VSVS: Grammaire latine: description du latin classique en vue de la lecture des auteurs*. Paris: Duculot.
- Lehmann, Thomas. 1993. *A grammar of modern Tamil*. Pondicherry: Pondicherry Institute of Linguistics and Culture.
- Lesko, Leonard H. & (in collaboration with) Barbara Switalski Lesko. 2002-2004. *A dictionary of late Egyptian*. 2nd edn. Providence: Scribe Publ.
- Leslau, Wolf. 1938. *Lexique soqotri (sudarabique moderne) avec comparaisons et explications étymologiques*. Paris: Klincksieck.
- Lewis, Charlton T., Charles Short, E. A. Andrews & William Freund. 1879. *A Latin dictionary*.
<http://www.perseus.tufts.edu/>.
- Lexa, František. 1947-1951. *Grammaire démotique*. Prague: Lexa.

- Li, Jinfang. 1996. Bugan - a new Mon-Khmer language of Yunnan province, China. *Mon-Khmer Studies* 26. 135-159.
- Li, Paul. 2006. The internal relationships of Formosan languages. *Tenth International Conference on Austronesian Linguistics, 17-20 January 2006*, Puerto Princesa City, Palawan, Philippines. <http://www.sil.org/asia/philippines/ical/papers.html>.
- Liddell, Henry George, Robert Scott, (revised by) Henry Stuart Jones & (with the assistance of) Roderick McKenzie. 1940. *A Greek-English lexicon*. Oxford: Clarendon Press. <http://www.perseus.tufts.edu>.
- Lieberman, Stephen J. 1990. Summary report: Linguistic change and reconstruction in the Afro-Asiatic languages. In Philip Baldi (ed.), *Linguistic change and reconstruction methodology*, 565-576. Berlin: Mouton de Gruyter.
- Lindström, Eva. 1995. Animacy in interrogative pronouns. In Inger Moen, Hanne Gram Simonsen & Helge Lødrup (eds.), *Papers from the XVth Scandinavian Conference on Linguistics (Oslo, January 13-15, 1995)*, 307-315. Oslo: University of Oslo.
- Linn, Mary S. 2000. *A grammar of Euchee (Yuchi)*: University of Kansas PhD thesis.
- Lipiński, Edward. 1997. *Semitic languages: Outline of a comparative grammar*. Leuven: Peeters.
- Lomtadidze, Ketevan V. 1967a. Абазинский язык [Abaza]. In Viktor V. Vinogradov *et al.* (ed.), *Языки народов СССР [Languages of the peoples of the USSR]*, Vol. 4: Иберийско-кавказские языки [Ibero-Caucasian languages], 123-144. Moscow: Nauka.
- Lomtadidze, Ketevan V. 1967b. Абхазский язык [Abkhaz]. In Viktor V. Vinogradov *et al.* (ed.), *Языки народов СССР [Languages of the peoples of the USSR]*, Vol. 4: Иберийско-кавказские языки [Ibero-Caucasian languages], 101-122. Moscow: Nauka.
- Loprieno, Antonio. 1996. *Ancient Egyptian: A linguistic introduction*. Cambridge: Cambridge University Press.
- Love, J.R.B. & (edited by) R.M.W. Dixon. 2000. *The grammatical structure of the Worora language of north-western Australia*. München: LINCOM Europa.
- Luce, G.H. 1965. Danaw, a dying Austroasiatic language. *Lingua* 14. 98-129.
- Macomber, William F. no date/1999? *A Bohairic-English dictionary*. <http://web.iasphost.com/stmark/coptic/coptdict.pdf>.
- Macuch, Rudolf. 1969. *Grammatik des samaritanischen Hebräisch*. Berlin: Walter de Gruyter & Co.
- Madvig, Johan Nicolai Agathon. 1870. *Grammaire latine*. Paris: Didot.
- Magomedbekova, Z. M. 2001. Ботлихский язык [Botlikh]. In Mikhail Ye. Alekseev, Georgy A. Klimov, Sergey A. Starostin & Yakov G. Testelefs (eds.), *Языки мира: Кавказские языки [Languages of the World: Caucasian languages]*, 228-236. Moscow: Academia.
- Maho, Jouni F. 2001. The Bantu area: (towards clearing up) a mess. In Jouni F. Maho (ed.), *Africa & Asia, 1 (Göteborg working papers on Asian and African languages and literatures)*, 40-49. Göteborg: Department of Oriental and African Languages, Göteborg University. www.african.gu.se/aa/pdfs/aa01040.pdf.

- Maho, Jouni F. 2002. *The Bantu line-up: comparative overview of three Bantu classifications*. Göteborg: Department of Oriental and African Languages, Göteborg University. <http://www.african.gu.se/maho/downloads/bantulineup.pdf>.
- Maia, Marcus A.R., Andrés Pablo Salanova & Elder José Lanes. 2000. La sintaxis de las interrogativas en karajá, kayapó y manxineri. In Hein van der Voort & Simon van de Kerke (eds.), *Ensaíos sobre lenguas indígenas de las tierras bajas de Sudamérica: contribuciones al 49o Congreso Internacional de Americanistas en Quito, 1997/ Essays on indigenous languages of lowland South America: contributions to the 49th International Congress of Americanists in Quito, 1997*, 297-307. Leiden: Universiteit Leiden, Research School of Asian, African, and Amerindian Studies.
- Malchic Nicolás, Manuel Bernardo, Romelia Mó Isém & Augusto Tul Rax. 2000. *Rub'iral ruk'ihhaal wach Poqom q'orb'al/ Variación dialectal en Poqom*. Guatemala: Cholsamaj.
- Mamet, M. 1960. *Le langage des Bolia (Lac Léopold II)*. Tervuren: Musée royal du Congo belge.
- Martins, Silvana Andrade. 2004. *Fonologia e gramática Dâw*. Amsterdam: Vrije Universiteit Amsterdam PhD thesis.
- Martins, Silvana Andrade & Valteir Martins. 1999. Makú. In R.M.W. Dixon & Alexandra Y. Aikhenvald (eds.), *The Amazonian languages*, 251-267. Cambridge: Cambridge University Press.
- Masica, Colin P. 1991. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.
- Maslova, Elena. 2000. A dynamic approach to the verification of distributional universals. *Linguistic Typology* 4(3). 307-333.
- Maslova, Elena. 2003. A case for implicational universals. *Linguistic Typology* 7(1). 101-108.
- Masqueray, Émile. 1879. Comparaison d'un vocabulaire des Zénaga avec les vocabulaires correspondants des dialectes Chawia et des Beni Mzab. *Archives des missions scientifiques et littéraires* 3(5). 473-533.
- Matson, Dan Mitchell. 1964. *A grammatical sketch of Juang, a Munda language*: The University of Wisconsin PhD thesis.
- Mayers, Marvin K. (ed.). 1966. *Languages of Guatemala*. The Hague: Mouton & Co.
- Maytinskaya, Klara Ye. 1969. Местоимения в языках разных систем [*Pronouns in languages of different types*]. Moscow: Nauka.
- Mažiulis, Vytautas. ms. [2004]. *Historical Grammar of Old Prussian*. (A translation of Lithuanian edition of 2004 and comments by Dr. L.Palmaitis). <http://donelaitis.vdu.lt/prussian/mazgr.pdf>.
- McConvell, Patrick. 2005. *The archaeo-linguistics of migration*. Paper presented at the University of Leuven, 13.12.2005.
- McGregor, William B. 1990. *A functional grammar of Gooniyandi*. Amsterdam: John Benjamins.
- McGregor, William B. 1994. *Warrwa*. München: LINCOM Europa.
- McGregor, William B. 1996. *Nyulnyul*. München: LINCOM Europa.
- McGregor, William B. 2004. *The languages of the Kimberley, Western Australia*. London: Routledge Curzon.

- McKay, G.R. 1975. *Rembarnga: A language of Central Arnhem Land*. Canberra: Australian National University PhD thesis.
- McLaughlin, John E. 2006. *Timbisha (Panamint)*. München: LINCOM Europa.
- Meeussen, Achilles E. 1967. Bantu grammatical reconstruction. *Africana linguistica* 3. 80-121.
- Mel'čuk, Igor A. 1993-2000. *Cours de morphologie générale (théorique et descriptive)*. Montréal: Presses de l'Université de Montréal.
- Meléndez Lozano, Miguel Ángel. 2000. Esbozo gramatical de la lengua achagua. In María Stella González de Pérez & María Luisa Rodríguez de Montes (eds.), *Lenguas indígenas de Colombia: una visión descriptiva*, 625-640. Bogotá: Instituto Caro y Cuervo.
- Merlan, Francesca. 1982. *Mangarayi*. Amsterdam: North-Holland Publishing Company.
- Merlan, Francesca. 1983. *Ngalakan grammar, texts and vocabulary*. Canberra: Pacific Linguistics.
- Merlan, Francesca. 1994. *A grammar of Wardaman: a language of the Northern Territory of Australia*. Berlin: Mouton de Gruyter.
- Messineo, Cristina. 2003. *Lengua toba (guayacurú): aspectos gramaticales y discursivos*. München: LINCOM Europa.
- Metzger, Ronald G. 1981. *Gramática popular del carapana*. Bogotá: Instituto Lingüístico de Verano.
- Militarev, Alexander Y. 2005. Once more about glottochronology and comparative method: the Omotic-Afrasian case, *Orientalia et Classica, Papers of the Institute of Oriental and Classical Studies*, Vol. 6: Аспекты компаративистики, 1 [Aspects of comparative linguistics, 1]. Festschrift for S.A. Starostin, 339-408. Moscow: Russian State University for the Humanities. <http://starling.rinet.ru/Texts/fleming.pdf>.
- Militarev, Alexander Y. & Olga V. Stolbova. no date. *Afro-Asiatic etymological database*. <http://starling.rinet.ru/cgi-bin/main.cgi?flags=eygtnnl>.
- Miller, Marion. 1999. *Desano grammar*. Dallas: Summer Institute of Linguistics.
- Miller, Vera G. 1976. *An overview of Stiêng grammar*. Grand Forks: Summer Institute of Linguistics, University of North Dakota.
- Milne, Leslie. 1921. *An elementary Palaung grammar*. Oxford: Clarendon.
- Mineau, R. 1982. *Les vieux parlers poitevins: histoire, phonétique, grammaire*. Poitiers: Brissaud.
- Missionários de Companhia de Jesus. 1964. *Elementos de gramática cinyanja*. Lisboa: Junta de investigações do ultramar.
- Mitchell, Bruce. 1985. *Old English syntax*. Oxford: Clarendon.
- Mithun, Marianne. 1999. *The languages of native North America*. Cambridge: Cambridge University Press.
- Mohapatra, Sasmita. 1991. *Juang [Grammar, texts, dictionary]*. Bhubaneswar: Academy of Tribal Dialects.
- Morse, Nancy L. & Michael B. Maxwell. 1999. *Cubeo grammar*. Dallas: Summer Institute of Linguistics.

- Mosonyi, Esteban Emilio. 2000. Breve caracterización conjunta de las lenguas curripaco y piapoco. In María Stella González de Pérez & María Luisa Rodríguez de Montes (eds.), *Lenguas indígenas de Colombia: una visión descriptiva*, 641-656. Bogotá: Instituto Caro y Cuervo.
- Mous, Maarten. 1993. *A grammar of Iraqw*. Hamburg: Helmut Buske.
- Müller, Heinrich. 1902. *Die Mehri- und Soqotri-Sprache*. Vol. 1: Texte. Wien: Hölder.
- Müller, Heinrich. 1905. *Die Mehri- und Soqotri-Sprache*. Vol. 2: Soqotri-Texte. Wien: Hölder.
- Müller, Heinrich. 1907. *Die Mehri- und Soqotri-Sprache*. Vol. 3: Šhauri-Texte. Wien: Hölder.
- Murkelinsky, Gadzhi B. 1967. Лакский язык [Lak]. In Viktor V. Vinogradov *et al.* (ed.), *Языки народов СССР [Languages of the peoples of the USSR]*, Vol. 4: Иберийско-кавказские языки [Ibero-Caucasian languages], 488-507. Moscow: Nauka.
- Mushin, Ilana. 1995. Epistememes in Australian languages. *Australian Journal of Linguistics* 15. 1-31.
- Mustanoja, Tauno F. 1960. *A Middle English syntax. Part 1: Parts of speech*. Helsinki: Société néophilologique.
- Nagaraja, K.S. 1985. *Khasi: A descriptive analysis*. Pune: Deccan College.
- Nagaraja, K.S. 1996. The status of Lyngngam. *Mon-Khmer Studies* 26. 37-50.
- Najlis, Elena Lidia. 1966. *Lengua abipona*. Buenos Aires: Centro de Estudios Lingüísticas, Universidad de Buenos Aires.
- Nash, David. 1997. Sketch of Warlmanpa by contrast with Warlpiri. *Language Tutorial, Conference of LAGB, Spring 1997*, University of Edinburgh.
<http://www.anu.edu.au/linguistics/nash/aust/wpa/wpa-wlp-contrast.html>.
- Natarajan, G.V. 1985. *Abujhmaria grammar*. Mysore: Central Institute of Indian Languages.
- Nau, Nicole. 1999. Was schlägt der Kasus? Zu Paradigmen und Formengebrauch von Interrogativpronomina. *Sprachtypologie und Universalienforschung* 52(2). 130-150.
- Naumkin, Vitaly V. & Viktor Ya. Porkhomovsky. 1981. *Очерки по этнолингвистике Сокотры [Ethnolinguistic studies of Soqotra]*. Moscow: Nauka.
- Nedjalkov, Igor. 1997. *Evenki*. London: Routledge.
- Nekes, Hermann, Ernest A. Worms & (edited by) William B. McGregor. 2006. *Australian languages*. Berlin: Mouton de Gruyter.
- Neumann-Holzschuh, I. 1987. *Textes anciens en créole louisianais*. Hamburg: Helmut Buske.
- Neumann, I. 1985. *Le créole de Breaux Bridge, Louisiane*. Hamburg: Helmut Buske.
- Nguyễn, Đình-Hoà. 1997. *Vietnamese*. Amsterdam: John Benjamins.
- Nichols, Johanna. 1988. Nominalization and assertion in scientific Russian prose. In John Haiman & Sandra A. Thompson (eds.), *Clause combining in grammar and discourse*, 399-428. Amsterdam: John Benjamins.
- Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.
- Nicolas, Francis. 1953. *La langue berbère de Mauritanie*. Dakar: Institut Français d'Afrique Noire.
- Nikolaeva, Irina & Maria Tolskaya. 2001. *A grammar of Udihe*. Berlin: Mouton de Gruyter.

- Nimuendajú, Curt. 1941 (ms.). *Witoto, Miranya und Kayuisána*. Rio de Janeiro: Arquivo da FUNAI.
- Nordlinger, Rachel. 1993. *A grammar of Wambaya*. Melbourne: University of Melbourne MA thesis.
- Nordlinger, Rachel. 1998. *A learner's guide to basic Wambaya*. Canberra: Australian Institute for Aboriginal and Torres Strait Islander Studies.
- Nyrop, Kristoffer & (edited by) Pierre Laurent. 1935. *Grammaire historique de la langue française*. Copenhagen: Gyldendalske Boghandel Nordisk Forlag.
- O'Grady, Geoffrey N. 1968. *Mirninny wordlist*. [Document 0070 of the Aboriginal Studies Electronic Data Archive, Australian Institute of Aboriginal and Torres Strait Islander Studies, Canberra, typed by David Nash, 1990. Extracted from: O'Grady 1968 comparative Australian material].
- O'Grady, Geoffrey N. 1979. Preliminaries to a Proto Nuclear Pama-Nyungan stem list. In Stephen A. Wurm (ed.), *Australian linguistic studies*, 107-139. Canberra: Pacific Linguistics.
- Oates, Lynette F. 1988. *The Muruwari language*. Canberra: Pacific Linguistics.
- Ogawa, Naoyoshi & Erin Asai. 1935. *Gengo ni yoru Taiwan takasagozoku densetsu-shû [Myths and traditions of the Formosan native tribes]*. Taipei: Taihoku Imperial University.
- Olawsky, Knut J. 2006. *A grammar of Urarina*. Berlin: Mouton de Gruyter.
- Onenko, Sulungu N. 1980. *Нанайско-русский словарь [Nanai-Russian dictionary]*. Moscow: Russkij jazyk.
- Ottaviano, John C. & Ida Ottaviano. 1967. Tacana. In Esther Matteson (ed.), *Bolivian Indian grammars*, Vol. 1, 139-208. Norman: Summer Institute of Linguistics of the University of Oklahoma.
- Par Sapón, María Beatriz & Telma Angelina Can Pixabaj. 2000. *Ujunamaxiik ri K'ichee' ch'ab'al/ Variación dialectal en K'ichee'*. Guatemala: Cholsamaj.
- Pardee, Dennis. 1997. Ugaritic. In Robert Hetzron (ed.), *The Semitic languages*, 131-144. London: Routledge.
- Pasch, Helma. 1995. *Kurzgrammatik des Ewe*. Köln: Rüdiger Köppe.
- Patal Majzul, Filiberto, Pedro Oscar García Matzar & Carmelina Espantazay Serech. 2000. *Rujunamaxik ri Kaqchikel chi' / Variación dialectal en Kaqchikel*. Guatemala: Cholsamaj.
- Patte, Marie-France. 2002. The Arawak language. In Eithne B. Carlin & Jacques Arends (eds.), *Atlas of the languages of Suriname*, 82-112. Leiden: KITLV Press.
- Patz, Elizabeth. 2002. *A grammar of the Kuku Yalanji language of north Queensland*. Canberra: Pacific Linguistics.
- Paulsen, Debby. 1992. A phonological reconstruction of Proto-Plang. *Mon-Khmer Studies* 18-19. 160-222.
- Payne, David L. 1980. *Diccionario ashéninka-castellano*. Yarinacocha: Instituto Lingüístico de Verano.

- Payne, David L. 1990. Some widespread grammatical forms in South American languages. In Doris L. Payne (ed.), *Amazonian linguistics: Studies in Lowland South American languages*, 75-87. Austin: University of Texas Press.
- Payne, David L., Judith K. Payne & Jorge Sánchez S. 1982. *Morfología, fonología y fonética del ashéninka del apurucayali (campa-arawak preandino)*. Lima: Ministerio de Educación & Instituto Lingüístico de Verano.
- Penchoen, Thomas G. 1973. *Tamazight of the Ayt Ndhir*. Los Angeles: Undena Publications.
- Pensalfini, Robert J. 1997. *Jingulu grammar, dictionary, and texts*. Cambridge: Massachusetts Institute of Technology PhD thesis.
- Pensalfini, Robert J. 2003. *A grammar of Jingulu, an aboriginal language of the Northern Territory*. Canberra: Pacific Linguistics.
- Pérez, Eduardo Gustavo, Zoila Blanca Luz García Jiménez & Odilio Jiménez. 2000. *Tx'ixpub'ente tiib' Qyool/ Variación dialectal en Mam*. Guatemala: Cholsamaj.
- Perkins, Revere D. 1989. Statistical techniques for determining language sample size. *Studies in Language* 13(2). 293-315.
- Pillai, R. Sreeveeramanikandan & P. Kothandaraman. 1972. Demonstratives and interrogatives in Tamil and Malayalam. In S. Agesthialingom & S.V. Shanmugam (eds.), *Third Seminar on Dravidian linguistics*, 347-360. Annamalainagar: Annamalai University.
- Pitman, Mary. 1981. *Diccionario araona y castellano*. Riberalta: Instituto Lingüístico de Verano & Ministerio de Educación y Cultura.
- Platts, John T. 1884. *A dictionary of Urdu, classical Hindi, and English*. London: W. H. Allen & Co. <http://dsal.uchicago.edu/dictionaries/platts/>.
- Plumley, Martin J. 1948. *Introductory Coptic grammar*. London: Home & Van Thal. <http://metalog.org/files/plumley/html/home.htm>.
- Plungian, Vladimir A. 2000. *Общая морфология: Введение в проблематику [General morphology: Introduction in the problematics]*. Moscow: Editorial URSS.
- Popovich, A. Harold. 1984. The sun and the moon, a Maxakali text. *Estudos sobre Línguas e Culturas Indígenas*. 29-59.
- Popovich, A. Harold. 1986. The nominal reference system of Maxakali. In Ursula Wiesemann (ed.), *Pronominal systems*, 351-358. Tübingen: Gunter Narr Verlag.
- Popovich, A. Harold & Frances B. Popovich. 2005. *Maxakali-English dictionary. English-Maxakali glossary*. Cuiabá: Summer Institute of Linguistics. <http://www.sil.org/americas/brasil/PUBLICNS/DICTGRAM/MXDicEN.pdf>.
- Porter, Doris. 1979. Northern Kankanay morphology. *Studies in Philippine Linguistics* 3(2). 20-62.
- Poulos, George. 1990. *A linguistic analysis of Venda*. Pretoria: Via Afrika.
- Poulos, George & Louis J. Louwrens. 1994. *A linguistic analysis of Northern Sotho*. Pretoria: Via Afrika.
- Prachakij-karacak, Phraya. 1995 [1919]. *Some languages of Siam*. Bangkok: Institute of Language and Culture for Rural Development, Mahidol University.
- Rabel-Heymann, Lili E. 1961. Gender in Khasi nouns. *Mon-Khmer Studies* 6. 247-272.

- Rabel, Lili E. 1961. *Khasi, a language of Assam*. Baton Rouge: Louisiana State University Press.
- Radetzky, Paula Kadose. 2004. Grammaticalisation d'un marqueur de définitude en saaroa. In Elizabeth S. Zeitoun (ed.), *Les langues austronésiennes*, 213-230. Paris: Ophrys.
- Rajabov, Ramazan N. 1997. The class category in Tsez: underlying principles. In Mikhail Ye. Alekseev (ed.), *Дагестанский лингвистический сборник: Памяти Г.А. Климова [Dagestan papers in linguistics: In memory of G.A. Klimov]*, Vol. 4, 127-137. Moscow: Academia.
- Ramamurti, G.V. 1931. *A manual of the So:ra: (or Savara) language*. Madras: The Superintendent, Government Press.
- Ramamurti, G.V. 1938. *Sora-English dictionary*. Madras. [Reprinted in 1986, Delhi: Mittal Publications]. <http://ling.ill.hawaii.edu/austroasiatic/AA/sora-english-dictionary.pdf>.
- Ramanarasimham, P. 1998. Old Telugu. In Sanford B. Steever (ed.), *The Dravidian languages*, 181-201. London: Routledge.
- Ramirez, Henri. 1997. *A fala tukano dos ye'pâ-masa*. Vol. 1: Gramática. Manaus: Inspetoria Salesiana Missionária da Amazônia, CEDEM.
- Ramirez, Henri. 2001. *Línguas Arawak da Amazônia Setentrional: Comparação e descrição*. Manaus: Editora da Universidade do Amazonas.
- Rangan, K., M. Suseela & S. Rajendran. 2002. Agreement system in Tamil. *Language in India* 1(10). <http://www.languageinindia.com/feb2002/rajendran2.html>.
- Ray, Sidney H. 1933. *A grammar of the Kiwai language, Fly Delta, Papua, with a Kiwai vocabulary by the late Rev. E. Baxter Riley*. Port Moresby: Edward George Baker, Government Printer.
- Raz, Shlomo. 1997. Tigré. In Robert Hetzron (ed.), *The Semitic languages*, 446-456. London: Routledge.
- Reddy, Joy. 1979. *Kuvi grammar*. Mysore: Central Institute of Indian Languages.
- Reed, Judy & David L. Payne. 1986. Asheninca (Campa) pronominals. In Ursula Wiesemann (ed.), *Pronominal systems*, 323-331. Tübingen: Gunter Narr Verlag.
- Reid, Lawrence A. (ed.). 1971. *Philippine minor languages: Wordlists and phonologies*. Honolulu: University of Hawaii Press.
- Reid, Lawrence A. 1979. Evidence for Proto-Philippine Nominative marking. *Philippine Journal of Linguistics* 10(1-2). 1-20.
- Reid, Lawrence A. 1981. Proto-Austronesian genitive determiners. In Andrew Gonzalez & David Thomas (eds.), *Linguistics across continents: Studies in honor of Richard S. Pittman*, 97-105. Manila: Summer Institute of Linguistics & Linguistic Society of the Philippines.
- Reid, Lawrence A. 1999. New linguistic evidence for the Austric hypothesis. In Elizabeth S. Zeitoun & Paul Jen-kuei Li (eds.), *Selected papers from the Eighth International Conference on Austronesian Linguistics*, 5-30. Taipei: Academia Sinica.

- Reid, Lawrence A. 2000. Sources of Proto-Oceanic initial prenasalization: The view from outside Oceanic. In Videa P. De Guzman & Byron Bender (eds.), *Grammatical analysis: Morphology, syntax, and semantics*, 30-45. Honolulu: University of Hawai'i Press.
- Reid, Lawrence A. 2002. Determiners, nouns or what? Problems in the analysis of some commonly occurring forms in Philippine languages. *Oceanic Linguistics* 41(2). 295-309.
- Reid, Lawrence A. 2006. On reconstructing the morphosyntax of Proto-Northern Luzon. *Tenth International Conference on Austronesian Linguistics, 17-20 January 2006*, Puerto Princessa City, Palawan, Philippines. <http://www.sil.org/asia/philippines/ical/papers.html>.
- Reid, Lawrence A. forthcoming. On reconstructing the morphosyntax of Proto-Northern Luzon. *Philippine Journal of Linguistics*.
- Reid, Lawrence A. & Hsiu-chuan Liao. 2004. A brief syntactic typology of Philippine languages. *Language and Linguistics* 5(2). 433-490.
- Reinhard, Johan & Sueyoshi Toba. 1970. *A preliminary linguistic analysis and vocabulary of the Kusunda language*. Kathmandu: Summer Institute of Linguistics & Tribhuvan University.
- Reinisch, Leo. 1878. Die Saho-Sprache. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 32. 415-464.
- Reinisch, Leo. 1889. *Texte der Saho-Sprache*. Wien: Alfred Hölder.
- Reinisch, Leo. 1890. *Wörterbuch der Saho-Sprache*. Wien: Alfred Hölder.
- Rey-Debove, Josette et al. 1996. *Le Nouveau Petit Robert: Dictionnaire alphabétique et analogique de la langue française*. Paris: Dictionnaires Le Robert.
- Riegel, Martin, Jean-Christophe Pellat & René Rioul. 2001. *Grammaire méthodique du français*. Paris: Quadrige.
- Rijkhoff, Jan & Dik Bakker. 1998. Language sampling. *Linguistic Typology* 2(3). 263-314.
- Rijkhoff, Jan, Dik Bakker, Kees Hengeveld & Peter Kahrel. 1993. A method of language sampling. *Studies in Language* 17(1). 169-203.
- Riley, E. Baxter & Sidney H. Ray. 1930. Sixteen vocabularies from the Fly River, Papua. *Anthropos* (25). 173-194, 831-850.
- Rodrigues, Aryon D. 1999a. Tupí. In R.M.W. Dixon & Alexandra Y. Aikhenvald (eds.), *The Amazonian languages*, 107-124. Cambridge: Cambridge University Press.
- Rodrigues, Aryon D. 1999b. Macro-Jê. In R.M.W. Dixon & Alexandra Y. Aikhenvald (eds.), *The Amazonian languages*, 165-206. Cambridge: Cambridge University Press.
- Rood, N. 1958. *Ngombe-Nederlands-Frans woordenboek (Dictionnaire ngombe-néerlandais-français)*. Tervuren: Musée royal du Congo belge.
- Roper, E.M. 1928. *Tu Beḍawie: An elementary handbook for the use of Sudan government officials*. Hertford: Stephen Austin and Sons.
- Rosch, E. 1978. Principles of categorization. In E. Rosch & B.L. Lloyd (eds.), *Cognition and categorization*, 27-48. Hillsdale, NJ: Lawrence Erlbaum.
- Ross, Malcolm. 2002. Jabêm. In John Lynch, Malcolm Ross & Terry Crowley (eds.), *The Oceanic languages*, 270-296. Richmond: Curzon.

- Rottet, Kevin J. 2004. Inanimate interrogatives and settlement patterns in Francophone Louisiana. *Journal of French Language Studies* 14(2). 169-188.
- Ruffolo, Roberta. 2005. *Topics in the morpho-syntax of Ibaloy, Northern Philippines*. Canberra: Australian National University PhD thesis.
- Ruhlen, Merritt. 1987. *A guide to the world's languages*. Vol. 1: Classification. London: Edward Arnold.
- Rumsey, Alan. 2000. Bunuba. In R.M.W. Dixon & Barry J. Blake (eds.), *Handbook of Australian languages*, Vol. 5, 34-152. Melbourne: Oxford University Press.
- Sagart, Laurent, Roger Blench & Alicia Sanchez-Mazas (eds.). 2005. *The peopling of East Asia: putting together archaeology, linguistics and genetics*. London: Routledge Curzon.
- Salzer, Meinke & Shirley Chapman. 1996. *Dicionário bilingue nas línguas Paumari e Portuguesa*. Porto Velho: Sociedade Internacional de Linguística.
- Samarin, William J. 1967. *A grammar of Sango*. The Hague: Mouton.
- Sandalo, Filomena. 1997. *A grammar of Kadiwéu: with special reference to the polysynthesis parameter*. Cambridge: Massachusetts Institute of Technology Press.
- Sands, Anna Kristina. 1989. *A grammar of Garadjari, Western Australia*. Canberra: Australian National University BA thesis.
- Sands, Bonny. 1998. *Eastern and Southern African Khoisan: Evaluating claims of distant linguistic relationships*. Köln: Rüdiger Köppe.
- Sands, Kristina. 1996. *The ergative in Proto-Australian*. München: LINCOM Europa.
- Sapir, Edward. 1930 [1992]. Southern Paiute, a Shoshonean language. In William Bright (ed.), *The collected works of Edward Sapir. X: Southern Paiute and Ute linguistics and ethnography*, 17-314.
- Sasse, Hans-Jürgen. 1976. Dasenech. In Lionel M. Bender (ed.), *The Non-Semitic languages of Ethiopia*, 196-221. East Lansing: Michigan State University.
- Sasse, Hans-Jürgen. 1981. Afroasiatisch. In Bernd Heine & Thilo C. Schadeberg (eds.), *Die Sprachen Afrikas*, 129-148. Hamburg: Helmut Buske.
- Saulwick, Adam. 2003. *A first dictionary of Rembarrnga*. Maningrida: Maningrida Arts & Culture.
- Schauer, Stanley & Junia Schauer. 2000. El yucuna. In María Stella González de Pérez & María Luisa Rodríguez de Montes (eds.), *Lenguas indígenas de Colombia: una visión descriptiva*, 515-532. Bogotá: Instituto Caro y Cuervo.
- Schebeck, Bernhard. 1974. *Texts on the social system of the Atinyamaṭaṭa people, with grammatical notes*. Canberra: Pacific Linguistics.
- Schmalstieg, William R. 1974. *An Old Prussian grammar*. University Park: Pennsylvania State University.
- Schmidt, P.W. 1919. *Die Personalpronomina in den australischen Sprachen*. Wien: Akademie der Wissenschaften in Wien.
- Schmidt, Wilhelm. 1906. Die Mon-Khmer-Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens. *Archiv für Anthropologie* 5. 59-109.
- Schulze, F.V.P. 1911. *A grammar of the Kuvi language, with copious examples*. Madras.

- Schürmann, Clamor W. 1844. *A vocabulary of the Parnkalla language spoken by the natives inhabiting the Western shores of Spencer's Gulf. To which is prefixed a collection of grammatical rules hitherto ascertained by C.W. Shürmann*. Adelaide: George Dehane.
- Schwartz, Linda J. & Timothy Dunnigan. 1986. Pronouns and pronominal categories in Southwestern Ojibwe. In Ursula Wiesemann (ed.), *Pronominal systems*, 285-322. Tübingen: Gunter Narr Verlag.
- Segerer, Guillaume. 2002. *La langue bijogo de Bubaque (Guinée Bissau)*. Leuven: Peeters.
- Segert, Stanislav. 1997. Phoenician and the Eastern Canaanite languages. In Robert Hetzron (ed.), *The Semitic languages*, 174-186. London: Routledge.
- Sharp, Janet C. 2004. *Nyangumarta: a language of the Pilbara region of Western Australia*. Canberra: Pacific Linguistics.
- Shaver, Harold. 1996. *Diccionario nomatsiguenga-castellano, castellano-nomatsiguenga*. Pucallpa: Ministerio de Educación & Instituto Lingüístico de Verano.
- Shaw, Mary & Helen Neuenswander. 1966. Achi. In Marvin K. Mayers (ed.), *Languages of Guatemala*, 15-48. The Hague: Mouton & Co.
- Shejxov, Enver M. 2004. *Сопоставительная грамматика лезгинского и русского языков. Морфология. Синтаксис [A contrastive grammar of Lezgi and Russian. Morphology. Syntax]*. Moscow: Academia.
- Shetler, Joanne & Michael Walrod. 1983. Semantic and thematic structure of discourse in Balangao. *Studies in Philippine Linguistics* 4(2). 10-87.
- Shiffman, Harold F. 1999. *A reference grammar of spoken Tamil*. Cambridge: Cambridge University Press. <http://ccat.sas.upenn.edu/plc/tamilweb/book.html>.
- Shoemaker, Jack S. & Nola K. Shoemaker. 1967. Essejja. In Esther Matteson (ed.), *Bolivian Indian grammars*, Vol. 1, 209-283. Norman: Summer Institute of Linguistics of the University of Oklahoma.
- Shorto, Harry L. 1963. The structural patterns of Northern Mon-Khmer languages. In Harry L. Shorto (ed.), *Linguistic comparison in South East Asia and the Pacific*, 45-61. London: School of Oriental and African Studies.
- Shorto, Harry L., (main editor) Paul Sidwell, (assisting editor) Doug Cooper & (assisting editor) Christian Bauer. 2006. *A Mon-Khmer comparative dictionary*. Canberra: Pacific Linguistics.
- Shvedova, N. Yu. et al. (ed.). 1980. *Русская грамматика [Russian grammar]*. Moscow: Nauka.
- Sidwell, Paul. 2002. Genetic classification of the Bahnaric languages: A comprehensive review. *Mon-Khmer Studies* 32. 1-24.
- Sidwell, Paul. 2005. Proto-Katuic phonology and the sub-grouping of Mon-Khmer languages. In Paul Sidwell (ed.), *SEALS XV: papers from the 15th meeting of the Southeast Asian Linguistics Society*, 193-204. Canberra: Pacific Linguistics.
- Sidwell, Paul. 2006. Preface. In (main editor) Paul Sidwell, (assisting editor) Doug Cooper & (assisting editor) Christian Bauer (eds.), *A Mon-Khmer comparative dictionary by Harry Shorto*, vii-xxv. Canberra: Pacific Linguistics.

- Sidwell, Paul & Pascale Jacq. 2003. *A handbook of comparative Bahnaric*. Vol. 1: West Bahnaric. Canberra: Pacific Linguistics.
- Siemund, Peter. 2001. Interrogative constructions. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language Typology and Language Universals*, 1010-1028. Berlin: de Gruyter.
- da Silva Facundes, Sidney. 2000. *The language of the Apurinã people of Brazil (Maipure/Arawak)*. Buffalo: University of New York PhD thesis.
- Simeone-Senelle, Marie-Claude. 1997. The Modern South Arabian languages. In Robert Hetzron (ed.), *The Semitic languages*, 378-423. London: Routledge.
- Simeone-Senelle, Marie-Claude. 2005. *Up to date assessment of the results of the research on the Dahalik language (December 1996 - December 2005)*.
<http://lacan.vjf.cnrs.fr/PDF/Publications/Senelle/DahlikBilan.pdf>.
- Simpson, Jane. 2002. *A learner's guide to Warumungu: Mirlamirlajinjiki Warumunguku apparrka*. Alice Springs: IAD Press.
- Simpson, Jane & Jeffrey Heath. 1982 (ms.). *Warumungu sketch grammar*.
- Simpson, Jane & Luise Anna Hercus. 1996. Wirangu, Western Desert and Thura-Yura. *Western Desert Workshop*, Australian Linguistic Institute, Canberra, July 1996.
- Snell, Betty E. 1998. *Pequeño diccionario machiguenga-castellano*. Lima: Instituto Lingüístico de Verano.
- Solntsev, Vadim M. & Tuệ Hoàng (eds.). 2001. *Материалы Российско-вьетнамской лингвистической экспедиции [Materials of the Russian-Vietnamese linguistic expedition]*. Vol. 4: Язык рук [The Ruc language]. Moscow: Vostochnaya literatura.
- Solntseva, Nina V. 1996. Case-marked pronouns in the Taoih language. *Mon-Khmer Studies* 26. 33-36.
- Solntseva, Nina V. & Vadim M. Solntsev. 2001. Морфология [Morphology]. In Vadim M. Solntsev & Hoàng Tuệ (eds.), *Материалы Российско-вьетнамской лингвистической экспедиции [Materials of the Russian-Vietnamese linguistic expedition]*, Vol. 4: Язык рук [The Ruc language]. Moscow: Vostochnaya literatura.
- Spiegelberg, Wilhelm. 1925. *Demotische Grammatik*. Heidelberg: Winter.
- Steever, Sanford B. 1998a. Gondi. In Sanford B. Steever (ed.), *The Dravidian languages*, 270-300. London: Routledge.
- Steever, Sanford B. 1998b. Malto. In Sanford B. Steever (ed.), *The Dravidian languages*, 359-387. London: Routledge.
- Steever, Sanford B. (ed.). 1998c. *The Dravidian languages*. London: Routledge.
- Steever, Sanford B. 1998d. Introduction to the Dravidian languages. In Sanford B. Steever (ed.), *The Dravidian languages*, 1-39. London: Routledge.
- Steindorff, Georg. 1979. *Koptische Grammatik mit Chrestomathie, Wörterverzeichnis und Literatur*. Hildesheim: Gerstenberg.
- Stokes, Bronwyn & William B. McGregor. 2003. Classification and subclassification of the Nyulnyulan languages. In Nicholas Evans (ed.), *The non-Pama-Nyungan languages of*

- northern Australia: comparative studies of the continent's most linguistically complex region*, 29-74. Canberra: Pacific Linguistics.
- Stoll, Otto. 1884. *Zur Ethnographie der Republik Guatemala*. Zürich: Orell Füssli und Co.
- Strong, James. 1890. *Hebrew Bible dictionary*. <http://www.crosswire.org/sword>.
- Stumme, Hans. 1899. *Handbuch des Schilhischen von Tazerwalt*. Leipzig: Hinrichs.
- Subbarao, Karumuri V. & Gracious Mary Temsen. 2003. Wh-question formation in Khasi. In Tej Ratna Kansakar, Mark Turin & (with assistance from) Nirmal Man Tuladhar (eds.), *Themes in Himalayan languages and linguistics*, 197-218. Heidelberg & Kathmandu: South Asia Institute & Tribhuvan University.
- Subrahmanyam, P.S. 1998. Kolami. In Sanford B. Steever (ed.), *The Dravidian languages*, 301-327. London: Routledge.
- Sunik, Orest P. 1968. Тунгусо-маньчжурские языки: Введение [Tungus-Manchu languages: Introduction]. In Viktor V. Vinogradov et al. (ed.), *Языки народов СССР [Languages of the peoples of the USSR]*, Vol. 5: Монгольские, тунгусо-маньчжурские и палеоазиатские языки [Mongolic, Tungus-Manchu and Paleoasiatic languages], 53-67. Leningrad: Nauka.
- Sweet, Henry. 1898. *A new English grammar: logical and historical*. Clarendon Press: Oxford.
- Talibov, B.B. 1976. *К истории изучения вопросов морфологии дагестанских языков [On the history of studies of the morphology of Dagestanian languages]*. Makhachkala.
- Taylor, Douglas. 1951a. Inflectional system of Island Carib. *International Journal of American Linguistics* 17(1). 23-31.
- Taylor, Douglas. 1951b. Morphophonemics of Island Carib (Central American dialect). *International Journal of American Linguistics* 17(4). 224-234.
- Taylor, Douglas. 1955. Phonemes of the Hopkins (British Honduras) dialect of Island Carib. *International Journal of American Linguistics* 21(3). 233-241.
- Taylor, Douglas. 1956a. Island Carib II: word-classes, affixes, nouns, and verbs. *International Journal of American Linguistics* 22(1). 1-44.
- Taylor, Douglas. 1956b. Island-Carib morphology III: locators and particles. *International Journal of American Linguistics* 22(2). 138-150.
- Taylor, Douglas. 1957. On the affiliation of "Island Carib". *International Journal of American Linguistics* 23(4). 297-302.
- Taylor, Douglas. 1958a. Island Carib morphology IV: syntactic notes, texts. *International Journal of American Linguistics* 24(1). 36-60.
- Taylor, Douglas. 1958b. The place of Island Carib within the Arawakan family. *International Journal of American Linguistics* 24(2). 153-156.
- Taylor, Douglas & Berend J. Hoff. 1980. The linguistic repertory of the Island-Carib in the seventeenth century: The men's language - a Carib pidgin? *International Journal of American Linguistics* 46(4). 301-312.
- Taylor, Gerald. 1993. Aladim ou Mil e uma noites na Amazônia: relato baniwa do Içana contado por Domingo de Souza Paiva em janeiro de 1985. *Amerindia* 18. 139-176.

- Teichelmann, Christian G. & Clamor W. Schürmann. 1840. *Outlines of a grammar, vocabulary, and phraseology of the Aboriginal language of South Australia, spoken by the natives in and for some distance around Adelaide*. Adelaide: The authors.
- Terrill, Angela. 1998. *Biri*. München: LINCOM Europa.
- Terrill, Angela. 2003. *A grammar of Lavukaleve*. Berlin: Mouton de Gruyter.
- Thomas, David D. 1971. *Chrau grammar*. Honolulu: University of Hawaii Press.
- Thornell, Christina. 1997. *The Sango language and its lexicon (Sëndâ-yângâ tí sängö)*. Lund: University of Lund.
- Till, Walter C. 1961. *Koptische Dialektgrammatik mit Lesestücken und Wörterbuch*. 2nd revised edn. München: Beck.
- Till, Walter C. 1986 [1966]. *Koptische Grammatik (Säidischer Dialekt) mit Bibliographie, Lesestücken und Wörterverzeichnissen*. 6th edn. Leipzig: Enzyklopädie Verlag.
- Titov, Evgenij G. 1991. *Граммати́ка амха́рского язы́ка [A grammar of Amharic]*. Moscow: Nauka.
- Trail, Ronald L. (ed.). 1973. *Patterns in clause, sentence, and discourse in selected languages of India and Nepal*. Vol. 4: Word lists. Kathmandu: Summer Institute of Linguistics & Institute of Nepal and Asiatic Studies, Tribhuvan University.
- Traill, Anthony. 1994. *A !Xóõ dictionary*. Köln: Rüdiger Köppe.
- Tryon, Darrell. 1970. *An introduction to Maranungku (Northern Australia)*. Canberra: Pacific Linguistics.
- Tsuchida, Shigeru. 1969 (ms.). *Saaroa word list*.
- Tsuchida, Shigeru. 1976. *Reconstruction of Proto-Tsouic phonology*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Tsunoda, Tasaku. 1995. *The Djaru language of Kimberley, Western Australia*. Canberra: Pacific Linguistics.
- Tung, T'ung-ho. 1964. *A descriptive study of the Tsou language, Formosa*. Taipei: Academia Sinica, Institute of Philology and History.
- Turner, Ralph Lilley. 1966. *A comparative dictionary of Indo-Aryan languages*. London: Oxford University Press.
- Ultan, Russell. 1978. Some general characteristics of interrogative systems. In Joseph H. Greenberg (ed.), *Universals of human language*, Vol. 4: Syntax, 211-248. Stanford: Stanford University Press.
- Valentine, J. Randolph. 2001. *Nishnaabemwin reference grammar*. Toronto: University of Toronto Press.
- Van de Velde, Mark. 1999. The two language maps of the Belgian Congo. *Annales Aequatoria* 20. 475-489.
- Van de Velde, Mark L.O. 2006. Multifunctional agreement patterns in Bantu and the possibility of genderless nouns. *Linguistic Typology* 10(2). 183-221.
- Van de Velde, Mark L.O. forthcoming. The alleged class 2a prefix *bò* in Eton, a plural word. In Rebecca Cover & Yuni Kim (eds.), *Proceedings of the Berkeley Linguistics Society 31st annual meeting*. Berkeley: University of California at Berkeley.

- Van Langendonck, Willy. 1999. Neurolinguistic and syntactic evidence for basic level meaning in proper names. *Functions of Language* 6. 95-138.
- Van Langendonck, Willy. 2007. *Theory and typology of proper names*. Berlin: Mouton de Gruyter.
- Vidal, Alejandra. 2001. *Pilagá grammar (Guaykuruan family, Argentina)*. Eugene: University of Oregon.
- Viñas Urquiza, María Teresa. 1974. *Lengua mataca*. Buenos Aires: Centro de Estudios Lingüísticos, Universidad de Buenos Aires.
- Vossen, Rainer. 1997. *Die Khoe-Sprachen: Ein Beitrag zur Erforschung der Sprachgeschichte Afrikas*. Köln: Rüdiger Köppe.
- Vycichl, Werner. 1984. *Dictionnaire étymologique de la langue copte*. Leuven: Peeters.
- Wallace, Judith M. 1966. Katu personal pronouns. *Mon-Khmer Studies* 2. 55-62.
- Watters, David E. 2006. Notes on Kusunda grammar (a language isolate of Nepal). *Himalayan Linguistics Archive* 3. 1-182.
http://www.uwm.edu/Dept/CIE/HimalayanLinguistics/Archive_2006/Watters.Kusunda.pdf.
- Wells, J. Warren. 2000-2006. *Sahidica: The Sahidic Coptic New Testament with parallel Greek*.
<http://www.integlogic.com/sahidica/sahidica.htm>.
- Werner, Heinrich. 1997. *Die ketische Sprache*. Wiesbaden: Harrassowitz Verlag.
- Wesley, John. 1754-1765. *Notes on the Bible*. <http://www.bible-center.ru>.
- Westermann, Diedrich. 1961. *Die Ewe-Sprache in Togo: Eine praktische Einführung*. Berlin: Walter de Gruyter & Co.
- Westphal, Ernst O.J. 1956. The non-Bantu languages of Southern Africa. In Archibald Norman Tucker & Margaret Arminel Bryan (eds.), *The non-Bantu languages of North-Eastern Africa*, 158-173. London: Oxford University Press.
- Westphal, Ernst O.J. 1962a. On classifying Bushman and Hottentot languages. *African Language Studies* 3. 30-48.
- Westphal, Ernst O.J. 1962b. A re-classification of Southern African non-Bantu languages. *Journal of African Languages* 1. 1-8.
- Westphal, Ernst O.J. no date (ms.). *Field notes on N|huki*. University of Cape Town, Manuscripts and archives department.
- Whaley, Lindsay. 1998. *Introduction to the study of Tungusic languages. Section 3: Basic typological features of Tungusic languages*. Hanover: Dartmouth College.
<http://www.dartmouth.edu/~trg/languages/section3.html>.
- Widmann, Thomas & Peter Bakker. 2006. Does sampling matter? A test in replicability, concerning numerals. *Linguistic Typology* 10(1). 83-95.
- Wierzbicka, Anna. 1996. *Semantics: Primes and universals*. Oxford: Oxford University Press.
- Wiesemann, Ursula. 1986. The pronoun systems of some Jê and Macro-Jê languages. In Ursula Wiesemann (ed.), *Pronominal systems*, 359-380. Tübingen: Gunter Narr Verlag.
- Wilkins, David P. 1989. *Mpartnwe Arrernte (Aranda): studies in the structure and semantics of grammar*. Canberra: Australian National University PhD thesis.

- Wilson, Peter J. & Stephen H. Levinsohn. 1992. *Una descripción preliminar de la gramática del achagua (Arawak)*. Bogotá: Instituto Lingüístico de Verano.
- Winfield, W.W. 1928. *A grammar of the Kui language*. Calcutta: The Asiatic Society of Bengal.
- Wise, Mary Ruth. 1986. Grammatical characteristics of preandine Arawakan languages of Peru. In Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*, Vol. 1, 567-642. Berlin: Mouton de Gruyter.
- Wogan, M. 1931. *Cancans Kisiniers*. New Orleans: Rogers Printing Co.
- Yen, Margaret *et al.* 1962-1963 (ms.). *Unpublished vocabularies of Saaroa and Kanakanabu taken by students of Prof. Tung T'ung-ho in 1962-1963*.
- Zaborski, Andrzej. 1976. Cushitic overview. In Lionel M. Bender (ed.), *The Non-Semitic languages of Ethiopia*, 67-84. East Lansing: Michigan State University.
- Zahn, H. 1940. *Lehrbuch der Jabémsprache (Deutsch-Neuguinea)*. Berlin: Verlag von Dietrich Reimer (Andrews & Steiner).
- Zamponi, Raoul. 2003. *Maipure*. München: LINCOM Europa.
- Zide, Norman H. 1969. Munda and Non-Munda Austroasiatic languages. In Murray B. Emeneau, Charles A. Ferguson & Gerald B. Kelly (eds.), *Linguistics in South Asia*, 411-430. The Hague: Mouton.

Nederlandse samenvatting

Een typologie van niet-selectieve interrogatieve pronomina

Deze lexicaal-typologische studie onderzoekt niet-selectieve interrogatieve pronomina die equivalent zijn aan Nederlands *wie?* en *wat?*. In het bijzonder bestudeer ik de diversiteit van natuurlijke (gesproken) talen op het gebied van de niet-selectieve interrogatieve pronomina ‘*wie?*’ en ‘*wat?*’ met betrekking tot: (i) de universaliteit van hun formele differentiatie en (ii) de patronen van hun functionele differentiatie in talen waar ze formeel worden onderscheiden. Dit onderzoek is ondernomen vanuit een functioneel-typologisch perspectief (bvb. zoals begrepen door Croft (2003)).

In de inleiding begin ik met het definiëren van het object van dit onderzoek, namelijk de niet-selectieve interrogatieve pronomina ‘*wie?*’ en ‘*wat?*’ (Sectie I.2). Daarna introduceer ik de onderzoeksvragen (Sectie I.3) en bespreek ik wat als een geval van het gebrek aan differentiatie tussen ‘*wie?*’ en ‘*wat?*’ beschouwd kan worden en wat niet (Sectie I.4). In Sectie I.5 geef ik een overzicht van het bestaande werk over deze onderwerpen. De algemene inleiding wordt afgesloten met een bespreking van de sample (Sectie I.6) en de manier waarop mijn data zijn verzameld (Sectie I.7).

Het centrale gedeelte van deze studie bestaat uit twee delen. Het eerste deel is gewijd aan een cross-linguïstisch onderzoek van patronen van functionele differentiatie tussen ‘*wie?*’ en ‘*wat?*’ in termen van niet-prototypische combinaties van waarden en ‘*wie?*’ of ‘*wat?*’-dominantie, zoals bepaald in Sectie I.4.2.3 en Sectie II.1. Na een aantal inleidende opmerkingen in Sectie II.1 bespreek ik eerst in Sectie II.2 het gebruik van ‘*wat?*’ in vragen over de classificatie van een persoon. Daarna, in Sectie II.3, bespreek ik het gebruik van ‘*wie?*’ in vragen over eigennamen van dingen en in Sectie II.4 het gebruik van ‘*wie?*’ in vragen over levende dingen. In Sectie II.5 wordt een conclusie gegeven.

In het tweede deel bespreek ik verschillende talen die lijken een gebrek aan differentiatie tussen ‘*wie?*’ en ‘*wat?*’ toe te laten. Ik heb deze talen in zeven geografische gebieden onderverdeeld: (i) Afrika en het Midden Oosten (Sectie III.2), (ii) Eurazië (Sectie III.3), (iii) Zuidoost Azië en Oceanië (Sectie III.4), (iv) Nieuw Guinea (Sectie III.5), (v) Australië (Sectie III.6), (vi) Noord Amerika (in het zuiden tot en met Panama, maar uitgezonderd de Caribische eilanden; Sectie III.7), (vii) Zuid Amerika (Sectie III.8). Binnen deze gebieden worden de talen genetisch gegroepeerd. Voor elke taal probeer ik eerst te bepalen of er werkelijk over een gebrek aan differentiatie tussen ‘*wie?*’ en ‘*wat?*’ gesproken kan worden. Zo ja, poog ik (voor zover de data en de tijd het toelaten) de origine van het gebrek aan differentiatie tussen ‘*wie?*’ en ‘*wat?*’ in elk specifiek geval te achterhalen.

ERRATA

MAJOR:

- p.22, footnote 23 should be as follows: “That is, if ‘who?’ and ‘what?’ must be distinguished in **predicatively** used non-selective interrogative pronominals, then also in **non-predicative** use”.
- p.562, example (2) and the sentence that follows it should be as follows:
 “(2) predicative > non-predicative
 That is, if ‘who?’ and ‘what?’ must be distinguished in **predicatively** used non-selective interrogative pronominals, then also in **non-predicative** use, whereas the opposite is not necessarily the case.”
- pp.16-17: (i) the question mark in (22c) should be deleted; (ii) the footnote 21 should be deleted, (iii) the footnote 22 should be “A definite description as an answer would normally be possible when accompanied by some additional explanation, such as *...the one we talked about yesterday*, or if it is a so-called “monoreferential appellative expression” (cf. Van Langendonck 2007:102-106), such as *the moon*”.
- p.26 the sentence before example (32) should be: “For instance, in Russian the non-selective interrogative pronominal *kto* ‘who?’ can be used not only for [person + identification], as in (32a), but also for [person + classification], as in (32b), ~~thus extending into the semantic domain prototypically covered by čto ‘what?’, at least on the parameter TYPE OF REFERENCE.~~”
- p.564, the sentence on the lines 2-4 should be: “Although I may have some ideas as regards the kind of entity that *x* belongs to, in principle I should not be expected to **express them**.”
- p.564, the second sentence of the second paragraph should be: “Indeed, if there is no need for the speaker to **express his/her hypotheses on the** kind of entity variable *x* belongs to, why should most languages bother their speakers with this issue by obliging them to distinguish between ‘who?’ and ‘what?’?”

MINOR:

- p.1, the first sentence should be: “This study in lexical typology investigates non-selective interrogatives **pronominals** equivalent to English *who?* (1) and *what?* (2).”
- p.1, the sentence after example (2) should be: “In particular, I explore the diversity of natural (spoken) languages in the domain of the non-selective interrogatives pronominals ‘who?’ and ‘what?’ as regards”
- p.4 the first sentence of the last paragraph should be: “...The hedge in the definition of pronominals specifying that a pronominal itself does not need to belong to the lexico-syntactic class of nouns is necessary to account for cases when questions about **the identity of** persons and things are expressed with conventionalized noun phrases or clausal constructions not based on nominal interrogative pronominals.”

- p.31, the last sentence of footnote 32 should be: “the referent at issue is **a thing** (or that the noun used to refer to **it** is not masculine)”
- p.57, example (20b) the gloss of the word *commandar* should be **commander** and not *what*.
- p.59 the 3rd sentence in the paragraph following example (26) should be: “In other words, this hierarchy implies that if, for instance, a given language allows for KIND-questions to be formulated **with** ‘what?’ when the question is about the functional affiliation of a person...”
- p.62, footnote 9 should be as follows: “It is also possible to ask the same question using *yàaya* ‘**how?**’ (Paul Newman, p.c.)...”
- p.82, footnote 34, the second sentence should be: “It should also be mentioned that only the names of months of Rapa Nui **origin** are proprial.”
- p.121: the number of the first example on this page should be **(2)** instead of (Y2)
- p.303, example (83b) the gloss of *tar* should be **that[NOM]** and that of *ekun* should be **IPW[NOM]**.