

On interrogative pro-verbs¹

Dmitry Idiatov & Johan van der Auwera
Center for Grammar, Cognition and Typology
University of Antwerp

1. Introduction

Certain interrogative pro-word types have been considered rare or even nonexistent (e.g., Katz & Postal 1964: 99; Weinrich 1963: 122; Ultan 1978; Donegan & Stampe 1983: 339; Zaefferer 1990: 227; Gil 2001; Hagège 2003). For instance, Katz & Postal (1964: 98) claim, in a rather *a priori* manner, that "...the range of constituents that can be 'questioned' is actually quite small and, with one possible exception, is restricted to Noun Phrase and probably to the Determiner constituent of NP". This excludes the possibility of interrogative pro-verbs, as illustrated in (1), and must therefore be rejected.

- (1) Chukchi (Chukotko-Kamchatkan; <http://privatewww.essex.ac.uk/~spena/Chukchee/chapter4.html>)
req-arkən-əm igirqej gənin ekək?
do.what?-PROGRESSIVE-EMPHASIS right.now 2SG-POSSESSIVE son.ABSOLUTIVE
'What is your son doing right now?'

Gil (2001) suggests that "question words can 'ask about' items belonging to major (or open) syntactic categories, but not minor (or closed) ones". This will allow for interrogative pro-verbs, but this is not very good either, for it fails to account for the fact that interrogative pro-verbs are a rare phenomenon. As to their rarity, Hagège (2003) advocates a principle of linguistic economy: "it is much more economical to split the questions into an interrogative word 'who?' or 'what?' + one of the two verbs with a generic meaning, i.e. 'be' and 'do'". This cannot be a full answer either. Why would a two word construction be more economical than a one word construction? And even if it is, language is as much prone to redundancy as to economy.

This paper aims to elucidate further the issue of the rarity of interrogative pro-verbs (section 3), against the background of a more general hypothesis on what constitutes a possible interrogative pro-word (section 2).

2. A possible interrogative pro-word

The general hypothesis contains two ingredients:

- (i) a statement on the generality of the presupposition going with constituent questions,
- (ii) a claim on the relevance of the endocentric vs. exocentric distinction.

2.1. Constituent questions & presupposition

All theories on constituent questions make room for the idea that a constituent question is based on a presupposition with a variable. In (2) the presupposition says that John saw somebody, and the variable is the 'somebody' whose identity the hearer is invited to disclose.

- (2) *Who did John see?*

As the example of an interrogative pro-verb illustrated in (1) with Chukchi shows, it is important to allow the predicate of the constituent question as a possible variable. In order to account for this, we propose the definition in (3).

- (3) A constituent question is a question that asks for an instantiation of the variable *x* in an *It is known that (possibly) HAPPEN/EXIST (... x ...)* structure. The structure is the presupposition that the

¹ This paper results from research conducted as a part of the GOA (Geconcerteerde Onderzoeksactie) "Mood and Modality" project of the University of Antwerp. We wish to thank the Research Council of the University of Antwerp for its financial support. We would also like to thank two anonymous reviewers for their comments on the abstract of this work submitted to the workshop "Syntax, Semantics, and Pragmatics of Questions" (9-13 August 2004, Université Henri Poincaré, France, a part of the 16th European Summer School in Logic, Language and Information).

situation under interrogation (possibly) exists, existed or will exist. The variable *x* constitutes the focus of the constituent question and of its possible answer, and it is formally expressed by an interrogative pro-word.

The element *possibly* appears in the presupposition of examples like (4). Its presence is necessary there, because of the “weak” existential nature of such a presupposition as compared to “stronger” existential presuppositions of examples like (5).

(4) *Who can solve this problem?*

(5) *Who has bitten into the apple?*

In (5) there is a clear existential presupposition that *there is* somebody who has bitten into the apple, whereas (4) 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 (6). Being rather a matter of pragmatics than semantics, rhetorical questions are not considered in this paper.

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

It is also worth mentioning that a presupposition can sometimes turn out to be false. This can be illustrated by the fact that questions (2), (4) and under certain conditions even (5)² can all be answered by *nobody*.

2.2. Endocentric vs. exocentric distinction

If we want to find out what categories of interrogative pro-words can exist in the languages of the world, we should first determine which elements can potentially be variables that value that of the HAPPEN/EXIST (...) predicate can depend on. Clearly, these are first of all phrasal categories and only indirectly terminal categories. In other words, every phrasal level element can potentially be an argument of the HAPPEN/EXIST (...) predicate. Hence, an interrogative pro-element for any phrase level element can theoretically exist: interrogative pro-noun phrase, interrogative pro-verb phrase, interrogative pro-adjective phrase, interrogative pro-adverb phrase, interrogative pro-numeral phrase, interrogative pro-adposition phrase, interrogative pro-complementizer phrase, etc.

However, in the lexicon we are not dealing with phrases, but with words which correspond to terminal categories in syntax. Therefore, in order to find out which interrogative pro-words can exist, a good use can be made here of a distinction between endocentric and exocentric constructions. Only endocentric constructions can be reduced to their heads, that is to the terminal categories that we are looking for, without being simply elliptical (cf., among others, Hartman & Stork 1972: 76, Crystal 1985: 109).³ Thus, it can be argued that interrogative pro-words can be of endocentric phrase creating categories only – in the traditional sense of ‘endocentricity’. Some examples are given in (7). The endocentricity parameter turns out to be somewhat difficult to apply to verb phrases in a straightforward way. Therefore, interrogative pro-verbs will be considered separately in section 3.

- | | | |
|-----|--|--|
| (7) | a. <i>Where do you live?</i> | ADVERB category is endocentric phrasal |
| | b. <i>In what town do you live?</i> | ADJECTIVE/ DEICTIC ATTRIBUTE category is endocentric phrasal |
| | c. <i>In what do you live?</i> | NOUN category is endocentric phrasal |
| | d. <i>*WH-ADPOSITION a town do you live?</i> | ADPOSITION category is exocentric phrasal |

Example (7b) needs some explanation because many people would say that *what* is a determiner there. First, *what* as found in (7b) can, at least in some contexts, be considered as an interrogative pro-adjective because it can be synonymous to *what kind of?* and the expected answer will then be something like (8).

(8) *I live in a BIG town.*

Second, a more natural interpretation of *what* in (7b) will however be to consider it as an adnominal deictic attribute. In what follows we will try to explain why the latter term is preferred to a more usual term of determiner.

Normally, under the label of determiner such different categories are subsumed as articles, adnominal demonstrative attributes, most adnominal indefinite attributes and possessives. There has been a certain trend in linguistics to consider determiners as the heads in the phrases [determiner + noun]. On the other

² For instance, if the speaker does not see the apple, but only knows that there is an apple in another room and s/he has only heard a sound resembling that of somebody biting into the apple, whereas in fact nobody has bitten into the apple at all.

³ In other words, endocentric constructions are constructions the head of which is the single grammatically obligatory element in the construction.

hand, a more traditional analysis of determiners is to treat them as dependents.⁴ This traditional analysis seems to us to be much more appropriate. In addition, it should be pointed out that according to Dryer (1992) the category of determiner is far from universal and, it should be added, far from homogeneous. As we see it, what are normally called “determiners” are better regarded as at least two different categories: articles and adnominal deictic attributes (specifiers). The latter category includes (at least in English) adnominal demonstrative attributes, most adnominal indefinite attributes and possessives. Adnominal deictic attributes (specifiers) constitute a phrase creating category, whereas articles do not. This is illustrated in (9). Consequently, only interrogative pro-adnominal deictic attributes are possible, but no interrogative pro-articles.

(9) [*exactly this*] *apple*, [*exactly my*] *apple* but not **[exactly the] apple*⁵

Moreover, in many languages articles and other so-called determiners can even cooccur. This is illustrated by example (10) from Italian.

(10) *la mia mamma* ‘(*the) my mommy’

There is another point worth mentioning here. Many languages show a strong preference for treating all possible kinds of their interrogative pro-adnominal attributes in one and the same way as to the categorial characterization of these elements. This becomes most clear in the case of languages with a rich morphology. For instance, in Russian all interrogative pro-adnominal attributes, such as *kakoj* ‘what, what kind of; which’, *kotoryj* ‘which (one)’, *chej* ‘whose’, are morphologically adjectives. As far as English is concerned, its interrogative pro-adnominal attributes, such as *what*, *whose* and *which*, seem to be all treated in the same way as the interrogative pro-word *which*. The latter is best considered as interrogative pro-adnominal deictic attribute and not as “interrogative determiner” as is normally done. This is especially true if we speak in terms of interrogative *pro*-forms and not just in terms of interrogative words. Consider (11), where the most natural answer will be (a) or possibly (b), less so (c) and least of all (d).⁶

- (11) – *Which apple?*
 (a) – *This/ that apple.*
 (b) – *His/ my apple.*
 (c) – *Some apple.*
 (d) – *??The apple.*

As is shown in (12), neither *which*, nor *what*, nor *whose* can normally cooccur with an article, a possessive, an adnominal demonstrative attribute, or an adnominal indefinite attribute. The word *exactly* is also used in (12) in order to further highlight combinatorial differences of the elements at issue.

- (12) (**the*) *what/ whose/ which* (**his/ *some*) *house* (*exactly*)?
 (a) – (**the/ *a/ exactly*) *this* (**his/ *some*) *house* (*exactly*)
 (b) – *a* (??*exactly*) *nice house* (**exactly*)
 (c) – (*exactly/ the*) *John’s house* (**exactly*)

All the facts just presented clearly show that *what* in (7b) is best considered as interrogative pro-adnominal deictic attribute and not as interrogative determiner. Moreover, the same holds true for such interrogative pro-words as *which* and *whose*.

3. Interrogative pro-verbs

3.1. Interrogative pro-predicates: interrogative pro-verbs vs. interrogative pro-“non-verbal predicates”

Interrogative pro-verbs, as illustrated in (1) are a rare phenomenon. However, based on a scrutiny of a more or less representative sample of some three hundred fifty languages,⁷ it is clear that some types of

⁴ See, for instance, Van Langendonck (2003: 176-177) for a more detailed account of the discussion on the head/dependent status of determiners.

⁵ In an appropriate context, the last variant can be possible, but only if the article is stressed. However, in this case, the article functions as an adnominal anaphoric demonstrative attribute rather than as an ordinary definite article. The form *the* here does not indicate just the definite, known status of the referent of the word *apple*, but directly refers to the previous occurrence of the word *apple* in the discourse.

⁶ The variant (d) is possible only in an appropriate context and if the article is stressed. However, in this case, the article functions as an adnominal anaphoric demonstrative attribute rather than as an ordinary definite article. The form *the* here does not indicate just the definite, known status of the referent of the word *apple*, but directly refers to the previous occurrence of the word *apple* in the discourse.

⁷ The sample consists of languages from more than 70 language families (including several isolates) plus a couple of pidgins and creoles. The language families were counted in accordance with the Ethnologue (www.ethnologue.com/family_index.asp). It should

what grammars call “interrogative verbs” are more frequent than others. The reason is that the sources consulted often use such label as “interrogative verbs” for elements that are best described as interrogative pro-“non-verbal predicates” (i.e. “nominal predicates”, “adjectival predicates”, etc.). An example of such an element is given in (13).

- (13) Tuvaluan (Eastern Malayo-Polynesian; Besnier 2000: 425)
Ne aa taulua olooga ki motu?
 NONPAST what your go+NOMINALIZER to islet
 ‘How did your trip to the islets go?’ (lit.: ‘What (was) your going to the islets’).

Interrogative pro-“non-verbal predicates” and interrogative pro-verbs can be unified in a superordinate category of interrogative pro-predicates.

As our sample has shown, interrogative pro-“non-verbal predicates” are more frequent than interrogative pro-verbs. In languages that do not use copulas they would seem to be completely normal even. It is interesting to note that there are even languages in which all or almost all interrogative pro-words regularly function as interrogative pro-“non-verbal predicates”. One such case is represented by Salishan languages, spoken on the coast of British Columbia, Canada, and in the adjacent region of the USA. According to Kroeber (1999: 247), “wh-questions in Salish generally take the form of clefts ... in which the wh-word evidently is not in a complementizerlike position, but rather is the predicate of the higher clause of a biclausal construction”; “the wh-word is the focus, and is followed by a relative clause introduced by an article” (Kroeber 1999: 262).

- (14) Thompson (Salishan; Kroeber 1999: 263)
swét k=wík-t-xʷ
 who ARTICLE=see-TRANSITIVE-2SG.TRANSITIVE.SUBJECT
 ‘Who did you see?’

3.2. Interrogative pro-verbs: types and frequency

It can be argued that the difference in frequency between the interrogative pro-“non-verbal predicates” and interrogative pro-verbs also hangs together with the centricity parameter. On the one hand, non-verbal predicates in a language without copulas are typically intransitive and therefore endocentric. Consequently, an interrogative pro-“non-verbal predicate” would be very natural in such a language. On the other hand, an ordinary VP can be either endocentric or exocentric, depending on whether the verb is intransitive or transitive, respectively. How can one then question a VP?

The first possibility is to construe the VP as an argument of a general non-interrogative verb (such as *do*), as in English.

- (15) a. – *What did he do?*
 b. – *He swam/ He killed a bird.*

A construction based on a general non-interrogative verb such as *do* + complement *what* can represent both an endocentric phrase and an exocentric phrase because the interrogative pro-noun *what* is categorized as endocentric phrase creating. This can be compared to the interrogative pro-adverb *where* that is also categorized as endocentric phrase creating and that can stand for both an endocentric *there*, which is an adverb as well, and an exocentric *in London*, which is an adposition phrase. The verb *do* in turn serves to indicate that the information asked about is a process or action and that it is expected to be coded by a verb. In other words, the use of the *do* + *what* construction successfully clears up the VP centricity dilemma via a “division of labour”.

The second possibility is to create a single interrogative pro-verb inviting the hearer to identify the process, as illustrated in (16). In this case, however, we get a rather weird situation where a construal that categorizes as endocentric phrase creating is expanded to stand for exocentric phrases of one and the same category and consequently to categorize as exocentric phrase creating at the same time.

- (16) Kayardild (a Tangkic language of Australia; Evans 1995: 371)
nyinka ngaaka-wath?
 2SG.NOMINATIVE what-INCHOATIVE
 ‘What are you doing?’

The third option is to use two interrogative pro-verbs, one intransitive ‘do what?’ and the other one transitive ‘do what to [sth/sb]?’ , as in Paamese.

be noted that the Ethnologue uses the label of language family not only for families in the traditional sense but also for linguistic phyla such as Niger-Congo and even Australian languages.

- (17) Paamese (Austronesian, Eastern Malayo-Polynesian; Crowley 1982: 159)
- a. *raise gosaa?*
rice do.what:3SG:REALIS
'How is it going with the rice?' (lit.: 'What is the rice doing?')
 - b. *ko-gosein tuu-mali?*
2SG-do.what.to:REALIS brother-2SG:MASCULINE
'What have you done to your brother?'

At first sight, the existence of transitive interrogative pro-verbs seems to contradict the centrality criterion, because these interrogative pro-verbs are categorized as exocentric phrase creating. This apparent controversy can be explained as follows.

First, it is important to notice that the intransitive interrogative pro-verb in this type can still be answered by both an answer like *He swam* and an answer like *He killed a bird*, the transitive interrogative pro-verb being nothing but a regular (applicative) extension of the intransitive one. The attested transitive interrogative pro-verbs always happen to be morphologically related to their intransitive counterparts and are never found as the only interrogative pro-verbs in a given language. Some regular transitivity morphology seems to be indispensable for a language to be able to acquire a transitive interrogative pro-verb.⁸ In other words, a transitive interrogative pro-verb is nothing but a luxury a language can perfectly do without. In the same way, many languages are happy with expressing their recipients, benefactors or goals with just an adpositional phrase in combination with a simple intransitive verb, whereas others can afford a special applicative affix as well.

Second, if one considers the centrality from a semantic perspective, a transitive interrogative pro-verb seems to be somewhat more acceptable or tolerable than, for instance, an interrogative pro-adposition.⁹ The reason is that a transitive verb is as a head semantically much more prominent and to a certain extent independent as compared to an adposition.¹⁰ As a result, the complement of a transitive verb can be considered as being modifierlike and, consequently, the whole verb phrase as being semantically rather endocentriclike. In order to demonstrate this, let us compare an adposition phrase like (*John lives*) *in London* and a verb phrase like (*They*) *eat apples*. As to the adposition phrase at issue, strictly speaking we cannot say whether the whole concept is about *in* or about *London*. Neither of them can be really interpreted as a hyponym to which the whole construction refers (in terms of Hudson 1990: 106), in the sense of *good friends* being a kind of subtype of friends. On the other hand, the verb phrase *eat apples* refers rather to a subtype of eating, just as *good friends* are a subtype of friends. This is also the reason why an *intransitive* interrogative pro-verb can be answered both by an answer like *He swam* and an answer like *He killed a bird*.

As far as the cross-linguistic frequency of the three options is concerned, the three strategies can be arranged as follows (in brackets approximate percentages based on our sample are given for each strategy):

- I. a general non-interrogative verb such as *do* + complement *what* ($\approx 90\%$)
- II. a single intransitive interrogative pro-verb 'do what?' ($\approx 7,5\%$)
- III. both an intransitive interrogative pro-verb 'do what?' and a transitive interrogative pro-verb 'do what to [sth/sb]?' ($\approx 2,5\%$)

Two other theoretically possible options should also be mentioned. The first one would be to create an intransitive interrogative pro-verb for intransitive situations only (an answer like *He swam*) and use Strategy I for transitive and intransitive situations. The other option would be to create only a transitive interrogative pro-verb and use Strategy I for intransitive and transitive situations. However, even though the centrality criterion does not rule out these options completely, it makes such languages extremely improbable, and, indeed, no such strategy has been reported yet. Why should this be so?

For a start, think about the following. Even if we simply consider these two combinations – Strategy I plus an intransitive interrogative pro-verb for intransitive situations only and Strategy I plus a transitive interrogative pro-verb for transitive situations only – as to their internal coherence, we will immediately

⁸ For instance, in our example (19b) from Paamese the transitive interrogative pro-verb *gosein* seems to include a reflex of an applicative-causative suffix *-ni* (cf. Crowley 1982: 151).

⁹ Provided, of course, an intransitive interrogative pro-verb is already available in the language.

¹⁰ Cf. in this respect Van Langendonck (2003: 173): "The head creates a semantic syntagmatic slot that allows for specific semantic fillers. This is particularly clear in the case of verbs, since the verb (often polysemous) establishes which semantic roles can go together with it... For instance, in the sentence *Peter bought us a drink*, the word *Peter* is the agent, *a drink* the patient, and *us* the beneficiary of the verb *bought*."

notice that they are both very redundant. Strategy I alone is absolutely sufficient, because it allows with equal ease for both an intransitive verb phrase as an answer and a transitive verb phrase as an answer.

In addition, the fact that such languages are highly improbable actually follows rather straightforwardly from what has just been said when explaining, first, why transitive interrogative pro-verbs happen to exist at all and, second, why an *intransitive* interrogative pro-verb can be answered by both an answer like *He swam* and an answer like *He killed a bird*.

Let us consider first the option using Strategy I plus a transitive interrogative pro-verb only. Being categorized as exocentric phrase creating, a transitive interrogative pro-verb already represents a certain violation of the centricity principle. Any violation of a general principle is naturally expected to be possible only provided some extra factors come into action. In other words, a transitive interrogative pro-verb seems to be able to appear in a language only provided some regular transitivizing morphology is available and a comparable endocentric lexeme (i.e. an intransitive interrogative pro-verb) already exists in the language. This is so, because of the “first things first” logic: if a language happens to have an interrogative pro-verb at all, then according to the centricity principle a language is naturally expected to first have an interrogative pro-verb which is endocentric phrase creating, that is an intransitive interrogative pro-verb. Even though this endocentric phrase creating verb can be answered both by an answer like *He swam* and an answer like *He killed a bird*, what matters is that the interrogative pro-verb itself is endocentric phrase creating. Only when such a compromise with the centricity principle is agreed upon, only then, provided some regular transitivizing morphology is available, a further deviation from the centricity principle can be made and a transitive interrogative pro-verb can come to existence.

As to the option using Strategy I plus an intransitive interrogative pro-verb for intransitive situations only (an answer like *He swam*), we have shown above that an intransitive interrogative pro-verb can as a rule be answered both by an answer like *He swam* and an answer like *He killed a bird* and we have also tried to explain why this should normally be the case. This implies that the option at issue is highly improbable. On the other hand, what we do have a good chance to find is the following. When an intransitive interrogative pro-verb appears in a given language, this language can still go on using the *do* + *what* construction as well. In other words, in one and the same language we will find Strategy I (i.e., the *do* + *what* construction) and an intransitive interrogative pro-verb which allows for *both* an intransitive verb phrase as an answer *and* a transitive verb phrase as an answer.

References

- Besnier, Niko (2000). *Tuvaluan: a Polynesian language of the Central Pacific*. London: Routledge.
- Crowley, Terry (1982). *The Paamese language of Vanuatu*. Canberra: ANU.
- Crystal, David (1985). *A dictionary of linguistics and phonetics*, 2nd edition. New York: Basil Blackwell.
- Donegan, P. J. & Stampe, D. (1983). Rhythm and the holistic organization of language structure. In: Richardson, J. F. et al. (eds.). *Papers from the Parasession on the interplay of phonology, morphology and syntax*. Chicago: Chicago Linguistic Society, pp. 337-353.
- Dryer, Matthew S. (1992). The Greenbergian word order correlations. In: *Language* 68, pp. 81-138.
- Evans, Nicholas D. (1995). *A grammar of Kayardild*. Berlin & New York: Mouton de Gruyter.
- Gil, David (2001). Discussion on the LingTyp List on “interrogative verbs + interrogative relators + indefinite ordinals”. Available at: <http://listserv.linguistlist.org/cgi-bin/wa?A2=ind0103D&L=lingtyp&P=R272>, <http://listserv.linguistlist.org/cgi-bin/wa?A2=ind0103D&L=lingtyp&P=R663>, <http://listserv.linguistlist.org/cgi-bin/wa?A2=ind0103E&L=lingtyp&P=R225>.
- Hartmann, R.R.K., & Stork, F.C. (1972). *Dictionary of language and linguistics*. London: Applied Science.
- Hagège, Claude (2003). *Whatted we to interrogative verbs*. An abstract for the 5th ALT Conference, Cagliari (Italy), 15-18 September 2003.
- Hudson, Richard (1990). *English Word Grammar*. Oxford: Blackwell.
- Katz, Jerrold J. & Postal, Paul M. (1964). *An Integrated Theory of Linguistic Descriptions*. Cambridge, Mass.: MIT Press.
- Krober, Paul D. (1999). *The Salish Language Family: Reconstructing Syntax*. Lincoln & London: University of Nebraska Press.
- Ulan, Russel (1978). Some General Characteristics of Interrogative Systems. In: Greenberg, Joseph H. et al. (eds.), *Universals of Human Language*, Vol. 4. Stanford: Stanford University Press, pp. 211-248.

- Van Langendonck, Willy (2003). Dependency. Basic Principles and Basic Issues. In: Ágel, Vilmos et al. (eds.). *Dependency and Valency: An International Handbook of Contemporary Research*, Vol. 1. Berlin & New York: Walter de Gruyter, pp. 170-187.
- Weinrich, Uriel (1963). On the Semantic Structure of Language. In: Greenberg, Joseph H. (ed.), *Universals of Language*. Cambridge, Mass.: MIT Press, pp. 114-171.
- Zaefferer, Dietmar (1990). On the coding of sentential modality. In: Bechert, Johannes & Bernini, Giuliano & Buridant, Claude (eds.), *Toward a Typology of European Languages*. Berlin, New York: Mouton de Gruyter, pp. 215-237.