

## **A Bantu path towards lack of differentiation between 'who?' and 'what?'**

Dmitry IDIATOV

### **Abstract**

This paper proposes a typologically informed hypothesis for the origins of the lack of differentiation between 'who?' and 'what?' in several Bantu languages of zone C. It is argued that the zone C general interrogatives 'who? / what?', as well as 'who?' interrogatives in numerous other Bantu languages commonly reconstructed as *\*n(d)a(n)i* 'who?', have developed out of a selective interrogative 'which one?' and ultimately a locative interrogative construction *\*[AG9(or AG7) COP CL16 'what?']* '(it) is where?'. Formally, the latter clause-level construction underwent univerbation and a subsequent formal reduction with occasional reinforcement by a postposed demonstrative root. An important conclusion is that no interrogative pronominal meaning 'who?' can be reconstructed for Proto Bantu.

## 1. Introduction<sup>1</sup>

Several Bantu languages of zone C allow for a lack of differentiation between the non-selective interrogative pronominals ‘who?’ and ‘what?’.<sup>2</sup> As far as I can judge from the data used in Bastin *et al.* (1999) and my additional data, which together cover more than 90% of the Bantu languages, zone C is the only zone with such languages.<sup>3</sup> Moreover, the lack of differentiation between ‘who?’ and ‘what?’ as found in zone C is also unattested elsewhere in Niger-Congo (cf. Idiatov 2007). Depending on the way one counts languages, the number of zone C languages allowing for a lack of differentiation between ‘who?’ and ‘what?’ varies from three to ten, or more. In (1), I provide a list of five languages with their respective ‘who?’ / ‘what?’ interrogatives that may be considered representative for the whole area. See also Map 1 for a visualization of their approximate locations.

- |     |    |          |  |
|-----|----|----------|--|
| (1) | a. | ndè / nê | Mbosi / Mboshi (some variants; C25; Bastin <i>et al.</i> ’s 1999 database) |
|     | b. | nâ       | Tetela (some variants; C71; Bastin <i>et al.</i> ’s 1999 database)         |
|     | c. | ná       | Mongo(-Nkundo) (C61; Hulstaert 1938, 1961, 1965, 1966, 1993)               |
|     | d. | ínɔ      | Ntomba-Inongo (C35a; Bastin <i>et al.</i> ’s 1999 database)                |
|     | e. | ńɔ       | Bolia (C35b; Mamet 1960:35)  |

A few cases of lack of differentiation between ‘who?’ and ‘what?’ in the languages of the world have been reported in the typological literature (e.g. Hjelmlev 1956; Maytinskaya 1969; Ultan 1978; Lindström 1995; Siemund 2001; Bhat 2004). See also Idiatov (2007) for a detailed cross-linguistic study and many more new cases.

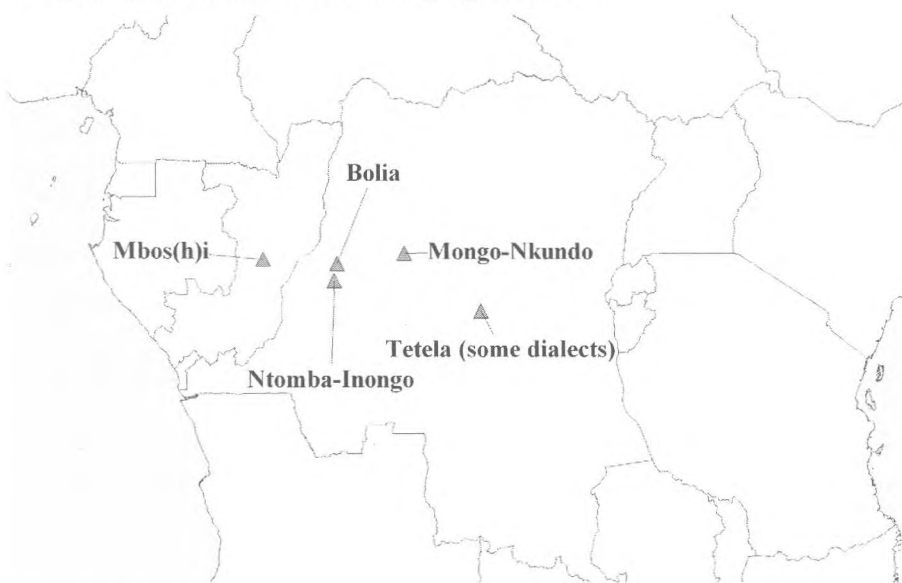
The data of the Bantu languages of zone C offer an insight in one of the ways through which languages can arrive at lack of differentiation between ‘who?’ and ‘what?’. The lack of differentiation between non-selective interrogative pronominals ‘who?’ and ‘what?’ in zone C receives a straightforward explanation when the typical origins of ‘who?’ interrogatives in Bantu are taken into consideration. Thus, in the Bantu languages where ‘who?’ is etymologically transparent, it can regularly be shown to go back to a selective interrogative pronominal meaning ‘which one? (person or thing)’, and ultimately to the locative interrogative ‘where?’. Similarly,

1. I am grateful to Yvonne Bastin, Koen Bostoen, Claire Grégoire, Johan van der Auwera and Mark Van de Velde for their helpful comments on the previous versions of this paper. Thanks are also due to anonymous referees for their constructive criticism. I also gratefully acknowledge the GOA (Geconcerteerde Onderzoeksactie) Mood and Modality Project of the University of Antwerp and the P6/44 project of the Interuniversity Attraction Poles programme of the Belgian Federal Government for financial support.  
 2. Traditionally, for reference purposes all Bantu languages are subdivided into 15-16 zones, named from A to S (with some gaps). Zones are further subdivided into groups and languages referred to by a two- digit number (sometimes expanded with a letter or a third digit). I use the codes from Guthrie (1967-1971) as updated by Maho (2009).  
 3. Bastin *et al.*’s (1999) database contains two languages outside of zone C marked as using a single interrogative ‘who?’ / ‘what?’, viz. Soli (M62) and a variety of Nyanja (N31a). 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 by class prefixes).

as I argue in Section 3, the lack of differentiation between ‘who?’ and ‘what?’ in several languages of zone C is due to the fact that their general interrogatives ‘who? / what?’ have developed out of a selective interrogative ‘which one?’ and ultimately a locative interrogative construction ‘(it) is where?’, originally indifferent to the distinction between persons and things, and that this indifference was maintained in these languages.

Reconstruction of short functional words is notoriously difficult for languages without long written traditions, as sound correspondences between functionally and formally comparable forms even at shallow time depths tend to be irregular. However, I believe a reconstruction may still be possible if we achieve a close match between (i) the observed variation in the reflexes of the element in question (also when the variation cannot be directly accounted for by regular sound correspondences), and (ii) typological knowledge of common processes of change that may provide us with some cues as to what historical sources may have produced the observed variation. I discuss the necessary evidence with regard to the reconstruction of the sources of the general interrogatives ‘who? / what?’ in Section 3.

Before proceeding to the diachronic discussion in Section 3, I will present the interrogative pronominal system of Mongo (C61) in Section 2 as an example of a zone C language allowing for a lack of differentiation between ‘who?’ and ‘what?’. Mongo is by far the best described language of zone C.<sup>4</sup>



Map 1. Some Bantu languages of zone C with ‘who?’ / ‘what?’ interrogatives.

4. It should be noted, though, that 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 (cf. Van de Velde 1999). 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 area around Boteka where the Flandria Mission of Hulstaert was located.

## 2. Mongo (C61)

Mongo, as described by Hulstaert (1938, 1961, 1965, 1966, 1993), has ná ‘who? / what?’ and é ‘what?’, as in (2-3).

Mongo (Hulstaert 1965:533)<sup>5</sup>

- (2) a. á-kela ná?  
AG1-do IPN  
‘What is he doing?’
- b. á-kela é?  
AG1-do IPN  
‘What is he doing?’
- (3) a. ilɔmbe yǎ ná?  
CL19.house AG19.CON IPN  
‘a house of whom?, whose house?’ or ‘a house of what?, a house for what (purpose, thing)?’
- b. ilɔmbe yǎ é?  
CL19.house AG19.CON IPN  
‘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),<sup>6</sup> 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 present nominal uses of ná and é in Section 2.1. Secondary uses will be considered in Section 2.2.

### 2.1. The interrogatives ná and é: nominal uses

Syntactically, ná ‘who? / what?’ and é ‘what?’ generally behave like nouns. They can be objects of verbs, as in (2) above, and of prepositions, as in (4).

Mongo (Hulstaert 1938:80)

- (4) a. lá ná?  
with IPN  
‘with whom?’ or ‘what for?, why?’

5. In Mongo examples, elisions are not represented and only classes and agreement patterns are regularly marked. Numbers of agreement patterns and classes follow the traditional Bantuist conventions, as found in Meeussen (1967).

6. 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 (3).

- b. lá      é?  
     with    IPN  
     'what for?, why?'

Based on the texts I analysed, é 'what?' is very rarely used. However, there is one context where it is much more common than ná, viz. in combination with the preposition lá 'with', as in (4), to ask 'what for?, why?'.<sup>7</sup> The words ná 'who? / what?' and é 'what?' can also function as a modifying element in a connective construction, as in (3). Moreover, they can be used as the second element of a nominal modifying construction without a connective,<sup>8</sup> as in (5), which can be compared to (6).

### Mongo

- (5) (lá)    ntsína      ná / é?  
       with   CL9.reason   IPN  
       'why?' (Hulstaert 1938:80)

- (6) a. baíso      jále  
       CL6.eyes   CL5.ferocity  
       'glaring eyes' (Hulstaert 1966:106)
- b. eyenga      bonkúnju  
       CL7.week   CL5.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's (have you seen yesterday)?* (e.g. John's house), *what about John (don't you like)?* or *who of the villagers (can afford it)?*, or (iii) heads of one of the few Mongo adjectives.<sup>9</sup> This implies that although ná and é are similar to nouns in other ways, they cannot occur in any position where they control agreement. Therefore, they should be best viewed as classless.

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, a pseudo-cleft construction has to be used. The predicate of a corresponding declarative clause is realized as a headless relative clause, while the interrogative is found in sentence-final position,

7. As suggested by a reviewer, lá é and lá ná may differ in their pragmatics. That is, lá é may be mostly for rhetorical questions, which would square well with the fact that é is also a polar question marker (cf. 2.2), whereas lá ná may be preferred when a real contenting answer is sought. This is a plausible hypothesis that is worth testing.

8. The term *connective* is used in Bantu studies to refer to a special element that is used to join two nominals in a construction where one nominal modifies the other in some way and that consists of a root on which agreement with the first nominal is marked by an agreement prefix. Connective in Bantu is largely similar to genitive in European languages.

9. 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.

as in (7). The relative can be analysed as the subject of the question and the interrogative as the predicate (cf. Hulstaert 1966:470).

Mongo (Hulstaert 1965:144)

- (7)      ḡ-kelaki              ná?  
          AG1.REL-did        IPN  
          ‘Who has done it? (lit.: ‘The one who has done (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 pseudo-cleft construction is also available for questioning objects, but it appears to be used less frequently than for questioning subjects.

Other zone C languages under consideration may impose less strict syntactic restrictions on their interrogative pronominals than Mongo does. For instance, in Bolia we find some sporadic examples of the interrogative pronominals questioning subjects in sentence-initial position, as in (8).

Bolia (C35b; Democratic Republic of the Congo; Mamet 1960:35, 125)

- (8)      ṅó              ḡ-tááká              itófbá              bentómba  
          IPN        AG1.REL-will.go        to.go.to.steal        cl.4.giant.pouched.rat  
  
          bē-lē              bó              ṅó?  
          AG4.REL-eat        PL        IPN  
          ‘Who would ever go to steal the giant pouched rats, who would eat them?’

Note, however, that the verb in (8) is still in the relative form and the interrogative is repeated sentence-finally, but now with the plural marker *bó*.

The use of the relative form in questions about subjects, as well as the sentence-final occurrence of interrogatives, is in accordance with 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). Relativization can be conceived of as a means of topicalizing the predicate, and indirectly, focalizing the subject. 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 contrastive focus, as in (9a),<sup>10</sup> the reverse order [relative form + notional subject] appears to be used for

10. The pre-verbal occurrence of the interrogative in the Bolia example (8) is probably due to the same reason. Note that the question in (8) 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 *bēlē bó nó* is best interpreted as a separate question: *Who (PL) would eat the rats?* Since the smoked rats

completive focus (Hulstaert 1966:472-473; see Dik 1989:282 for the notions of contrastive and completive focus), as in (9b). The latter order is also used in answers on questions about subjects (Hulstaert 1966:471).

Mongo (Hulstaert 1966:472)

- (9) a. ô nyama ð-wéi  
just CL9.animal AG1.REL-died  
‘It is just an ANIMAL that died (and not a man).’
- b. ð-wéi ô nyama  
AG1.REL-died just CL9.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 (9), 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 (7) to (10).

Mongo

- (10) [Bisénga, betúbyaíso!]  
Wolo nkína témpelo žyéngwéyá wolo,  
CL3.gold or CL9.temple AG1.REL-sanctify CL3.gold  
ð-lekí ná?  
AG1.REL-be.superior IPN  
‘[Ye fools and blind:] for whether is greater, the gold, or the altar that sanctifieth the gold?’ (Matthew 23:17 in Hulstaert 1987)<sup>11</sup>

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 preceded by a plural proclitic baa, as in (11).

Mongo (Hulstaert 1965:144)

- (11) bă-kelákí baa ná?  
AG2.REL-do PL IPN  
‘Who has done it? (lit.: ‘The ones who did it are who?’)’

## 2.2. The interrogatives ná and é: other uses

As was indicated in the beginning of Section 2, the interrogatives ná and é can also have some non-nominal uses. Thus, é can function as a polar question marker, as in (12).

---

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 (10) is selective, viz. ‘which one?’. Although in principle, this interpretation is possible, it is preferable to interpret ná in (10) as ‘what?’, given that Mongo has a separate word for ‘which one?’, viz. -lénkó.

Mongo (Hulstaert 1965:533)

- (12)    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 semantically rather straightforward and typologically well attested. 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 (13).

Mongo (Hulstaert 1965:533)

- (13)    ǎ-otswá       é?  
           AG1-went       where  
           ‘Where has he gone?’

As suggested by a reviewer, the use of é as ‘where?’ with verbs of movement originates in all probability in its use as a polar question marker. This development would involve conventionalization of a conversational implicature of the kind ‘Has he gone away?’ → ‘Where has he gone?’.

The form ná also functions as an exclamative postnominal modifier, as in (14).

Mongo (Hulstaert 1966:167)

- (14)    lifokú       ná!  
           CL5.beauty   IPN  
           ‘What beauty!’

This usage can be considered to be an instance of a nominal modifying construction without a connective, as in example (6). 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.<sup>12</sup> 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 nominally as ‘who?’ and as a nominal modifier ‘what [N]?, what kind of [N]?’ (that does not agree in class with the nominal) is attested elsewhere in Bantu, even in languages where ‘who?’ and ‘what?’ are otherwise different. Consider, for instance, the interrogative mang from Tswana (S31) in (15a-b), which otherwise is used as ‘who?’, as in (15c).

12. A reviewer suggested that (14) should be analysed rather as an N N nominal predicative construction, literally ‘beauty (is) which?’, reinterpreted as ‘what beauty!’. As such, this would be a possible analysis. However, it implies a semantic shift from ‘which one is beauty?’ to ‘what beauty!’, which semantically is much less straightforward (and typologically unusual) than from ‘what beauty?’ to ‘what beauty!’, where only the type of speech act changes, viz. from question to exclamation.



## Tswana (S31)

- (15) a. *thipa mang?*  
 CL9.knife IPN  
 ‘what (kind of) knife [do you like]?’ (Andy Chebanne, p.c.)
- b. *ke nako mang?*  
 COP CL9.time IPN  
 ‘What time is it?’ ([http://en.wikipedia.org/wiki/Tswana\\_language](http://en.wikipedia.org/wiki/Tswana_language))
- c. *molaodi ke mang?*  
 CL1.commander COP IPN  
 ‘[A:] Who is the commander? [B: It’s John / the tall man over there]’ (Andy Chebanne, p.c.)

The use of *mang* in (15a-b) can be compared to that of *-fe* ‘which [N]?; which one?’ in (16), which unlike *mang*, agrees in class with the noun it modifies, as many other nominal modifiers, such as adjectives, regularly do in Tswana.

## Tswana (S31)

- (16) *thipa e-fe?*  
 CL9.knife AG9-IPN  
 ‘which knife [did you take]?’ (Andy Chebanne, p.c.)

### 3. Origins of the lack of differentiation between ‘who?’ and ‘what?’ in zone C

The Mongo interrogative *é* ‘what?’ arguably goes back to the Proto Bantu interrogative *\*kì-í*, where *\*kì-* is the CL7 marker and the interrogative root *\*-í* can safely be reconstructed with the meaning ‘what?’ (or, in combination with one of the three Proto Bantu locative class prefixes, ‘where?’ and ‘which one?’; cf. Meeussen 1967:103; *BLR* 3; Doneux 1971:131-138; Doneux & Grégoire 1977:186-187; Schadeberg 2003:163).

The situation with the Mongo interrogative *ná* ‘who? / what?’ is less clear. Judging from Meeussen (1967:103), Doneux & Grégoire (1977:193), Schadeberg (2003:163) and the database *BLR* 3, *ná* and similar forms in other Bantu languages are reflexes of the following reconstructed forms: *\*n(d)ai* (Schadeberg 2003:163, tones uncertain) / *\*n(d)á-í* (the other sources), *\*ndá-ní* or *\*(n)á-ní*, neither of which is reconstructable (with certainty) to the Proto Bantu level. All the aforementioned sources suggest reconstructing these forms with the meaning ‘who?’, apparently due to the fact that it is the meaning that most reflexes of these forms have. 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, such as those mentioned in the introduction in (1). A shift from ‘who?’ to ‘who? / what?’ cannot be accounted for by any regular mechanisms of semantic change.<sup>13</sup> Moreover,

13. In this respect, it is worth emphasizing that such an alleged shift cannot be justified by appealing to the oft-cited examples of lexical-semantic shifts from animate (or human) to inanimate (or non-human) such as the expansion of English *body* to uses such as *car body*. Unlike the alleged shift from ‘who?’ to ‘who? / what?’, the latter kind of lexical-semantic shifts have a straightforward explanation as metaphoric extensions based on similarity in form, function or other properties of the referents of the respective nominals.

such a shift would be highly unusual typologically (cf. Idiatov 2007:574-575).

I argue that *\*n(d)a(n)i* should be reconstructed rather as a selective interrogative meaning 'which one? (person or thing)'. This reconstruction squares well with both (i) the fact that in many languages the reflexes mean 'who?' and (ii) the fact that in a few languages, such as Mongo, the reflexes can be used as both 'who?' and 'what?'. I further argue that the ultimate source of this selective interrogative is a locative phrase '(it) is where?'.

A historical link from 'which one?' to 'who?' (rather than 'what?') is not uncommon cross-linguistically (Idiatov 2007) and is pervasive in Bantu. Thus, the evolution from 'which one?' to 'who?' appears to repeat itself again in various Bantu languages. For instance, Doneux (1971) provides an illuminating discussion of the ongoing replacement in most Bantu languages of zone J of the older form for 'who?' by a new one from 'which one?'. Interestingly, the latter itself is clearly derived from 'where?' constructed as the combination of the CL16 prefix and the interrogative root *\*-í* 'what?'. For instance, in Ziba (JE22d) we find *hai* 'where?', *AG-hai* 'which (one)?', *ba-hai* 'who? (PL)', and *owai* 'who? (SG)' (Doneux 1971:126). In fact, 'which (one)?' recurrently turns out to be based on 'where?' in Bantu (cf. Doneux & Grégoire 1977:191).<sup>14</sup> Note that although the human semantics of Ziba *ba-hai* 'who? (PL)', *owai* 'who? (SG)' and similar forms in other Bantu languages is largely tied to the respective class prefixes, this does not undermine the validity of the claim that there is a preferential link from 'which one?' to 'who?' rather than 'what?' in Bantu. Thus the selective root 'which one?' has a non-selective reading only with a prefix of the so-called human AG1 / CL1 (and 2, when the intended meaning is plural), i.e. as 'who?', whereas with the prefixes of other classes and agreement patterns only the selective reading 'which one? (thing)' remains the norm rather than the non-selective 'what?'. Furthermore, in various Bantu languages where the interrogatives meaning 'who?' appear to represent a reflex of the locative interrogative *\*pà-í(-ní)*, such as Makwa-Ile (P31) *pání* and Giryama (E72a) *hani*,<sup>15</sup> no overt human AG1 / CL1 prefix shows up that can be held responsible for the human meaning 'who?'.

The development from 'which one?' to a general interrogative used as both 'who?' and 'what?' 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 being maintained. Apparently, this selective interrogative carried no overt human AG1 / CL1 prefix. Note in this respect that it is common for the interrogatives 'who?' in Bantu, especially those presumably representing reflexes of *\*n(d)a(n)i*, to lack any overt human AG1 / CL1 prefix. In fact, it is exactly due to this that these forms are typically set apart (together with several other nominals) as a special subclass of the human CL1 as the so-called CL1a (cf. Van de Velde 2006).

14. For instance, in Kagulu (G10/G12) we find *hoki* 'where?', *AG-hoki* 'which (one)?' and *ye-hoki* 'who? (SG)', *we-hoki* 'who? (PL)' (Petzell 2008:90, 177). Recall also the Mongo interrogative *AG-lénkó* 'which one?' (< AG-COP 'where?').

15. As is suggested by the similarity of their initial parts to the CL16 prefix *\*pà-* and the fact that in quite a few other languages similar interrogatives mean 'where?' and are clearly analysable as containing the CL16 prefix.

The locative-selective hypothesis of the origin of \*n(d)a(n)i can also help us to account in a straightforward way for (i) the modifying use (with both human and non-human nouns) of its reflexes that also mean ‘who?’, as was presented in Section 2.2, and for (ii) the use of its reflexes meaning ‘who?’ in questions about proper names (both personal and non-personal ones), as in (17) from Tswana and (18a) from Ngombe.

Tswana (S31; Andy Chebanne, p.c.)

- (17) a. lefelo je le bidiwa mang?  
 CL5.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?  
 CL5.name AG5.CON CL7.tree AG7.this COP who  
 ‘What is the name of this tree? (lit.: ‘The name of this tree is who?’)’ (e.g. oak)

Ngombe (C414; Rood 1958:xxi)<sup>16</sup>

- (18) a. ngando íyě nda?  
 CL9.village AG9.this who  
 ‘What is the name of this village? (lit.: ‘Who is this village?’)’
- b. moto íyö nda?  
 CL1.man AG1.this who  
 ‘Who is this man?’

In Mongo the situation is similar (at least for personal proper names), as can be seen in (19), where ná cannot be replaced with é ‘what?’ and therefore is glossed as ‘who?’.

Mongo

- (19) lína líkě ná / \*é?  
 CL5.name AG5.2SG.POSS who? / \*what?  
 ‘What is your name?’ (Honoré Vinck, p.c.)

As far as questions about proper names are concerned, it is not uncommon cross-linguistically for ‘where?’ and especially ‘which one?’ (rather than ‘what?’ or ‘who?’) to be used in questions about proper names (Idiatov 2007:62–63). If we assume that this was also the case in the Bantu languages at issue, this would mean that when the shift from ‘which one?’ to ‘who?’ occurred, these languages must have simply kept on using the same interrogative in questions about proper names.<sup>17</sup>

The locative-selective hypothesis proposed for \*n(d)a(n)i is also plausible from a formal point of view. In order to demonstrate this, let us consider the structure of these forms in more detail. To begin with, given that (i) at least in \*n(d)ai the

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

17. Bantu languages are probably not unique in this respect. For instance, a similar evolution appears to have occurred in some Cushitic languages (Idiatov 2007:73–74).

final element is apparently the Proto Bantu interrogative root *\*-í* ‘what?’<sup>18</sup> and that (ii) the ‘element n(d)a- is not attested otherwise’ (cf. Meeussen 1967:103; Doneux 1971:127; Doneux & Grégoire 1977:187, 191) and, therefore, cannot be directly attributed any lexical meaning of its own, it is not unreasonable to suppose that *\*n(d)a-* is further decomposable. Thus the vowel *a* could derive from the CL16 marker *\*pà-*, so that the ending *ai* would represent Proto Bantu *\*pà-í* ‘where?’ (cf. Meeussen 1967:103; Doneux 1971:133-138; Doneux & Grégoire 1977:187-191). Admittedly, in certain languages the loss of *\*p-* in *\*pà-* here is somewhat irregular but it can be accounted for by univerbation with the initial part *n(d)-*. Before proceeding to the discussion of the possible origins of the initial part *n(d)-*, let us consider the final element *-ni* in *\*n(d)a(n)i*.

The final element *-ni* may have two origins. First, it may correspond to the interrogative root *\*-ní* ‘which?’ reconstructed by Meeussen (1967:107). Doneux & Grégoire (1977:186-188) and Grégoire (1975:126) also provide quite a few examples of languages from zones A, B, C, H, K and L where locative interrogatives involving the CL16 marker *\*pà-* are based on the root *-ni* rather than *-i*, as in e.g. Tsogo (B31) *va-ni* ‘where?’, Geiriku (K332) *pa-ni* ‘where?’, Bobangi (C32) *wa-ni* ‘where?’, or Kele (C55) *á-ni-ma* ‘where?’.<sup>19</sup> It is also possible that *\*-ní* ‘which?’ itself goes back to a combination of some demonstrative root (*\*-na*, *\*-ni* or the like) and *\*-í* ‘what?’, so that originally the literal meaning of *\*-ní* was something like ‘what one?’.

The second possibility is that *-ani* derives from *\*pà-í-ní*, where *\*-ní* is some kind of reinforcing element, most likely of demonstrative origin,<sup>20</sup> added in some languages to reflexes of the locative interrogative *\*pà-í*. I prefer this hypothesis to the first one. To begin with, this hypothesis allows us to dispense with the additional interrogative root *\*-ní* ‘which?’. Its reconstruction for the Proto Bantu level is rather questionable anyway. At the same time, the reinforcement of a locative interrogative would not be unnatural for Bantu. Thus, according to Doneux (1971:134-135), in many languages of zone J the interrogatives meaning ‘where?’ have been reinforced through (i) reduplication or (ii) addition of initial *nka-* or final *-na*. Also recall the Kele (C55) interrogative *ánima* ‘where?’ mentioned above. Furthermore, the possibility of *-ani* going back to *\*pà-í-ní* is supported by forms such as Duala (A24) *we(ni)* ‘where?’, Punu (B43) *ave(ni)* ‘where?’ (Doneux & Grégoire 1977:187-188), where (i) *-ni* is an optional element and (ii) the vowel preceding it is *e* rather than *a*, which can be explained as the coalescence of the vowels of *\*pà-í*. In the latter respect, note also the variation between *a* and *e* in forms such as Luba (L31) *panyì / penyì* ‘where?’ (Kabuta *et al.* 2006).<sup>21</sup> In fact, it is quite common in Bantu for the sequence *ai* to become simplified as *a*, *e* or *i* (Doneux & Grégoire 1977).

18. This is so due to (i) the fact that the sequence *ai* is otherwise extremely rare in the reconstructed Proto Bantu morphemes (cf. Doneux & Grégoire 1977), and (ii) the interrogative function of *\*n(d)ai* that matches that of *\*-í*.

19. The final *-ma* in *ánima* is the noun *áma* ‘place’ (cf. Grégoire 1975:124).

20. For instance, the demonstrative stem *-ni* is widespread in zone C (Claire Grégoire, p.c.), as in Bolia *-níí* ‘that (very far, out of sight)’ (Mamet 1960:31).

21. The palatalization of *n* to *ny* in this context in Luba appears to be regular. For instance, compare the loanword *màshinyì* ‘car; (sewing) machine’ from French *machine* (Kabuta *et al.* 2006).

The initial *n(d)-* in *\*n(d)a(n)i* is likely to derive from the copula (or predicative marker) *\*ndí / ndí* ‘(it) is’, similar to the Venda (S21) invariable copula *ndí* (Poulos 1990:369), its Swahili (G42) analogues *ni* and *ndi-AG-o / ndi-* short form of *PRN* (as well as many so-called affirmative (nominal) predicative markers in various other Bantu languages) or the Mongo preposed focus particle *nd(é) / é* (Hulstaert 1965:638–639).<sup>22</sup> That is, the locative interrogative *\*pà-í* was introduced by *\*ndí / ndí* ‘(it) is’ in a kind of focalizing cleft construction ‘(it) is where (that...)?’. The vowel of the copula can also help to account straightforwardly for the palatalization/affrication and subsequent assibilation of *d* in the reflexes of *\*nda(n)i* in various languages from different zones, including many languages in zone C, such as Boloki *ndzai / njái* (C36e), Dzeke Babole (C101) *ndza* ‘who?’, Zamba (C322) *ńzai / nzai* (Bastin *et al.*’s 1999 database).

It is easy to conceive how in certain contexts the reconstructed predicative interrogative ‘(it) is where?’ could also be used with a selective meaning ‘(it) is which one? (person or thing)’ and how in the course of time it became confined to such (non-locative) selective uses. In turn, the shift of this interrogative to a non-locative meaning accompanied by its univerbation (i.e. fusion of a syntactic construction into a single word) obscuring its predicative origins facilitated its further spread to other typically nominal non-predicative contexts.<sup>23</sup> It is difficult to say how its ultimate expansion to non-selective contexts correlates with this syntactic evolution in time.

It is worth pointing out that the reconstruction *\*n(d)a(n)i* appears to miss one important element, viz. an initial *\*í-*. First, the reconstruction of the initial *\*í-* is suggested by forms such as Tsogo (B31) *índa* ‘who?’, Pinji (B304) *inde* ‘who?’ or Hunde (D51) *inde* (SG), *bende / bande* (PL) ‘who?’ (Bastin *et al.*’s 1999 database). Second, the original presence of an initial high-tone *\*í-* helps to account for the syllabic *ń-* in some of the interrogatives presented in (1), such as Ntomba-Inongo *ńń* and Bolia *ń*,<sup>24</sup> and for the presence of a floating high tone before all interrogative pro-words in Bolia (cf. Mamet 1960:34). This initial *\*í-* is unlikely to be an innovation given the attested scope of its areal distribution and the fact that it cannot be attributed any specific meaning or function in the respective languages. The possible function of this initial *\*í-* in the proto-form is suggested by the reconstructed copulative origin of *n(d)-* in *\*n(d)a(n)i*. Thus *\*í-* can be brought back to the subject prefix of AG9 *\*(j)í-* or perhaps AG7 *\*kí-*.<sup>25</sup> It is quite common for these agreement patterns to be used in Bantu as the default

22. That in Mongo *d* was preserved in the focus particle *nd(é) / é* but not in the interrogative *ná* is likely to be accounted for by the divergent development of the two forms and univerbation involved in the evolution of the interrogative, which would not be an uncommon historical process (e.g. compare English *gonna* and *going to*, *a* and *one*, etc.).

23. In this respect, cf. also Güldemann (2003:188) who advocates a comparable reanalysis of ‘short clauses’ as ‘non-predicative pronominal words’ in the history of Swahili (G42).

24. Interestingly, in Mongo we also find a similar form *ńa* as a kind of sentence-initial polar question marker. This *ńa* may have developed out of the same source as *ná* ‘who? / what?’, in which case *ńa* must be the older form. The relation between *ńa* and *ná* may be compared to that between *one* and *a(n)* in English.

25. The locative AG24 *\*í-*, reconstructed by Meeussen (1967), would also fit here. However, its reconstruction is not universally accepted.

agreement patterns when agreement is enforced (Van de Velde 2006:203-204),<sup>26</sup> i.e. when an atypical controller such as an interjection is used or no controller is present at all, as for instance in the case of an (agreeing) copula in a cleft sentence. Consider (20) from Eton (A71) for AG7 and (21) from Akwa (C22) for AG9.

Eton (A71)

- (20) í-nê            ùkèŋ            wó            újífbân  
AG7-COP    CL3.knife    AG3.PRN    AG3.PST.be.stolen  
'It is a knife that is stolen' (Van de Velde 2008:323)

Akwa (C22)

- (21) e-di            angwe  
AG9-COP    CL2.liana  
'[A: What is this? B:] These are lianas' (lit.: 'It is lianas') (Aksenova & Toporova 2002:266)

In many Bantu languages, the interrogatives meaning 'who?' that contain a reflex of the locative interrogative \*pà-í(-ní), such as pani, ha(n)i, fiani, ani, wani, weni, (j / y)wa(n)i, etc., do not show any traces of univerbation with a preceding copula. A case by case reconstruction of all these interrogatives falls outside of the scope of the present study.

Finally, a few words need to be said about the vowel quality in forms such as Ntomba-Inongo (C35a) níŋ and Bolia (C35b) níŋ presented in (1). In these languages ɔ is not a regular correspondence to Mongo a and unlike e in Mboshi (C25) ndè / nê (1), it cannot be explained as a coalescence of earlier a and i either. This ɔ may result from \*\*pa- > \*wa- > ɔ. In this respect, consider Bolia njóŋ / njǎ 'snake' corresponding to Mongo njwá (Proto Bantu \*-jóká), as well as Bobangi (C32) wa-ni 'where?', where, as mentioned earlier in the present section, wa- < \*pa-. This ɔ may also result from a merger of earlier a with the common Bantu demonstrative-cum-pronominal stem \*-o, sometimes also referred to as reference marker.<sup>27</sup> This is plausible phonologically since in Bolia a sometimes merges with a following o into ɔ (cf. Mamet 1960:17).<sup>28</sup> The latter hypothesis may be supported by a parallel from Kagulu (G12). Thus, Petzell (2008:90) argues that Kagulu hoki 'where?'<sup>29</sup> 'most likely contains the remnants of class 16 ha- and conceivably also the reference marker -o', with -ki being an 'interrogative clitic' meaning 'what?'.

#### 4. Conclusion

In the present paper, I have argued that the 'who? / what?' interrogatives attested in several Bantu languages of zone C originate in a locative interrogative construction

26. The term *enforced agreement* is attributed to Corbett (1991:204), who notes that in a gender language, if an agreement target *can* agree then typically it *must* agree.

27. Cf. the Proto Bantu anaphoric pronominals \*AG-o and demonstratives \*VOWEL-AG-o meaning 'that (near you); that (mentioned)' (Meeussen 1967:105-107).

28. The same happens in Mongo (cf. Hulstaert 1961:18).

29. Also AG-hoki 'which (one)?' and ye-hoki 'who? (sg)', we-hoki 'who? (pl)'.



meaning ‘(it) is where?’. This construction was first expanded to selective questions as ‘(it) is which one? (person or thing)’. The latter selective meaning gradually ousted its original locative meaning. The shift of this interrogative to a non-locative meaning accompanied by its univerbation obscuring its predicative origins facilitated its further spread to other typically nominal non-predicative contexts. In turn, the selective interrogative ‘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?’.<sup>30</sup>

On the formal side, the original construction underwent univerbation and a subsequent formal reduction. This often resulted in somewhat irregular sound changes, which is quite typical for these processes as is well-known from the history of the languages for which written sources of sufficient time-depth are available. In the case of Bantu, the reality of these changes can be ascertained thanks to the availability of numerous modern forms often presenting only minor variations.<sup>31</sup> Taken together with the known ways of semantic change of interrogative prowords, as are common in Bantu and crosslinguistically, this allowed me to argue that the zone C ‘who?’ / ‘what?’ interrogatives, such as Mongo ná, and similar interrogative pronominals with an initial n(d)- in other Bantu languages that are commonly reconstructed as something like \*n(d)a(n)i ‘who?’ should be brought back to the structure \*[AG9(or AG7)-COP CL16-‘what?’] ‘(it) is where?’, viz. something like \*ĩ-ndĩ / ndĩ pà-ĩ, which later fused into n(d)ai and the like. In many other languages, this structure was reinforced at a certain point by a postposed demonstrative root, most probably something like \*ni (or \*na), resulting in forms such as n(d)ani. Reflexes of the univerbated form are widespread in Bantu. However, there is also an important number of languages that lack such reflexes. This may be due either to the fact that no univerbation has ever occurred in these languages or that the older univerbated form has been completely ousted by a new form of a similar selective and ultimately locative origin (the process we can observe going on in many languages of zone J, for instance).

An important consequence of the present discussion is that no interrogative pronominal meaning ‘who?’ can be reconstructed for Proto-Bantu. Indeed, the only form that has been proposed in the literature for this role, viz. \*n(d)a(n)i (Meeussen 1967:103; Doneux & Grégoire 1977:193; Schadeberg 2003:163; *BLR* 3), cannot be reconstructed

30. A few languages outside of zone C demonstrate a situation intermediate between these two. In these languages, the form functioning on its own as ‘who?’ also serves as a base for ‘what?’. In this respect, consider the Luba (L31) interrogatives ɽanyì ‘who?’ and ci-ɽanyì ‘what?’ (Kabuta *et al.* 2006) and the Mwani (G403) interrogatives nani ‘who?’ and ki-nani ‘what?’ (with the CL7 prefix ki-) (Bastin *et al.*’s 1999 database).

31. An alternative solution dispensing with the possibility of such irregular changes due to univerbation and chunking would require us to reconstruct lots of different but accidentally very similar proto-forms for a linguistic group of a rather shallow time-depth such as Bantu. As such, this cautious approach would be methodologically sound. However, given that it is rather unlikely that such numerous similar reconstructions would reflect the historical reality, when compelling evidence is presented it is reasonable to prefer an alternative analysis as the best hypothesis (cf. Ricquier & Bostoen 2008 on ‘osculance’ in Bantu), even if it involves a certain degree of irregularity in sound correspondences. Of course, it is desirable that such irregularity can itself be accounted for by some well attested processes of linguistic change (such as univerbation and chunking).

as 'who?', but goes back to the locative interrogative 'where?' reinforced with some demonstrative element and/or unverbated with a predication element '(it) is'.

The Bantu facts discussed in the present paper provide an example of the nonuniversal nature of the differentiation between 'who?' and 'what?', show how languages can develop a general 'who?/what?' interrogative and provide an example of interrogative pronominals derived through unverbation of a clause-level construction.

### Glossing conventions

AG	agreement pattern
CL	noun class
CON	connective
COP	copula
IPN	interrogative pronominal
PL	plural
POSS	possessive
PRN	personal pronominal
PQ	polar question
PST	past
REL	relative
SG	singular

### References

- BLR 3 = Bastin, Y. & Schadeberg, T.C. (eds). 2005. *Bantu Lexical Reconstructions* 3. Tervuren: Royal Museum for Central Africa. <http://www.metafro.be/blr>.
- Aksenova, I.S. & Toporova, I.N. 2002. *Грамматика языка аква [A grammar of Akwa]*. Moscow: Editorial URSS.
- Bastin, Y., Coupeuz, A. & Mann, M. 1999. *Continuity and divergence in the Bantu languages: perspectives from a lexicostatistic study*. Tervuren: Royal Museum for Central Africa.
- Bhat, D.N.S. 2004. *Pronouns: A cross-linguistic study*. Oxford: Oxford University Press.
- Corbett, G.G. 1991. *Gender*. Cambridge: Cambridge University Press.
- Dik, S. 1989. *The theory of functional grammar*. Vol. 1: The structure of the clause. Dordrecht: Foris publications.
- Doneux, J.L. 1971. Deux études sur les interrogatifs en zone J du bantou. *Africana linguistica* V: 119-138.
- Doneux, J.L. & Grégoire, Cl. 1977. 'Une séquence vocalique à la finale de quelques thèmes en protobantou'. *Africana linguistica* VII: 159-199.
- Grégoire, Cl. 1975. *Les locatifs en bantou*. Tervuren: Royal Museum for Central Africa.
- Guthrie, M. 1967-1971. *Comparative Bantu*. London: Gregg Press.



- Güldemann, T. 2003. 'Grammaticalization'. In D. Nurse & G. Philippson (eds), *The Bantu languages*, 182-194. London: Routledge.
- Hjelmslev, L. 1956 [1971]. 'Animé et inanimé, personnel et non-personnel'. In L. Hjelmslev (ed.), *Essais linguistiques*, 211-249. Paris: Les Éditions de Minuit.
- Hulstaert, G. 1938. *Praktische grammatica van het Lonkundo (Lomongo) (Belgisch Kongo)*. Antwerpen: De Sikkel.
- 1961. *Grammaire du lomongo*. Vol. 1: La phonologie. Tervuren: Royal Museum for Central Africa.
- 1965. *Grammaire du lomongo*. Vol. 2: La morphologie. Tervuren: Royal Museum for Central Africa.
- 1966. *Grammaire du lomongo*. Vol. 3: La syntaxe. Tervuren: Royal Museum for Central Africa.
- 1987. *Bonkándá wă Njakomba (Elaká ěa Yóóko)/Le Nouveau Testament en lomongo*. Mbandaka: Imprimatur Mbandaka.
- 1993. 'Les interrogatifs dans les dialectes môngo'. *Annales Aequatoria* 14: 345-377.
- Idiatov, D. 2007. *A typology of non-selective interrogative pronominals*. Antwerp: University of Antwerp PhD.
- Kabuta, N.S. et al. (2006). *Dictionnaire cilubà-français*. <http://www.ciyem.ugent.be>.
- Lindström, E. 1995. 'Animacy in interrogative pronouns'. In I. Moen, H.G. Simonsen & H. Lødrup (eds), *Papers from the XVth Scandinavian Conference on Linguistics (Oslo, January 13-15, 1995)*, 307-315. Oslo: University of Oslo.
- Maho, J.F. 2009. NUGL online: The online version of the New Updated Guthrie List, a referential classification of the Bantu languages. June 4, 2009 edn. <http://goto.glocalnet.net/mahopapers/nuglonline.pdf>.
- Mamet, M. 1960. *Le langage des Bolia (Lac Léopold II)*. Tervuren: Musée royal du Congo belge.
- Maytinskaya, K.Y. 1969. *Местоимения в языках разных систем [Pronouns in languages of different types]*. Moscow: Nauka.
- Meeussen, A.E. 1967. 'Bantu grammatical reconstruction'. *Africana linguistica* III: 80-121.
- Missionários de Companhia de Jesus. 1964. *Elementos de gramática cinyanja*. Lisboa: Junta de investigações do ultramar.
- Petzell, M. 2008. *The Kagulu language of Tanzania: Grammar, texts and vocabulary*. Köln: Rüdiger Köppe.
- Poulos, G. 1990. *A linguistic analysis of Venda*. Pretoria: Via Afrika.
- Ricquier, B. & Bostoen, K. 2008. 'Resolving phonological variability in Bantu lexical reconstructions: the case of 'to bake in ashes''. *Africana linguistica* 14: 109-149.
- Rood, N. 1958. *Ngombe-Nederlands-Frans woordenboek (Dictionnaire ngombe-néerlandais-français)*. Tervuren: Musée royal du Congo belge.
- Schadeberg, T.C. 2003. 'Historical linguistics'. In D. Nurse & G. Philippson (eds), *The Bantu languages*, 143-163. London & New York: Routledge.
- Siemund, P. 2001. 'Interrogative constructions'. In M. Haspelmath, E. König, W. Oesterreicher & W. Raible (eds.), *Language Typology and Language*

- Universals*, 1010-1028. Berlin: de Gruyter.
- Ulan, R. 1978. 'Some general characteristics of interrogative systems'. In J. H. Greenberg (ed.), *Universals of human language*, Vol. 4: Syntax, 211-248. Stanford: Stanford University Press.
- Van de Velde, M.L.O. 1999. 'The two language maps of the Belgian Congo'. *Annales Aequatoria* 20: 475-489.
- 2006. 'Multifunctional agreement patterns in Bantu and the possibility of genderless nouns'. *Linguistic Typology* 10(2): 183-221.
- 2008. *A grammar of Eton*. Berlin: Mouton de Gruyter.

### Adresse de l'auteur

Dmitry IDIATOV  
 University of Antwerp  
 Center for Grammar, Cognition and Typology  
 Prinsstraat 13  
 2000 Antwerpen  
 Belgium  
 dmitry.idiatov@ua.ac.be

### Résumé

L'article formule, dans une perspective typologique, une hypothèse pour expliquer l'absence de différenciation entre 'qui ?' et 'quoi ?' dans de nombreuses langues bantoues de la zone C. Il soutient que les interrogatifs généraux 'qui ?/quoi ?', en zone C, ainsi que 'qui ?' dans de nombreuses autres langues bantoues – généralement reconstruits comme \*n(d)a(n)i 'qui ?' – se sont développés à partir d'un interrogatif sélectif signifiant 'lequel ?' et finalement à partir d'une construction locative à sens interrogatif \*[AG9(or AG7) COP CL16 'quoi ?'] '(cela) est où ?' Sur le plan formel, cette dernière construction a subi par la suite un processus d'univerbation ainsi qu'une réduction formelle avec occasionnellement un renforcement par un thème démonstratif postposé. Une conclusion importante de l'article est que, en proto-bantou, il n'est pas possible de reconstruire un pronom interrogatif signifiant 'qui ?'.