

Jenneke van der Wal* and Jacky Maniacky

How ‘person’ got into focus: Grammaticalization of clefts in Lingala and Kikongo areas

Abstract: In several Bantu languages in the regions where Kikongo and Lingala are spoken, we encounter sentences where the word ‘person’ can appear after the subject of a canonical SVO sentence, resulting in a focused interpretation of the subject. Synchronically, we analyze this as a monoclausal focus construction with *moto* ‘person’ as a focus marker. Diachronically, we argue, the construction derives from a biclausal cleft, where *moto* functioned as the head noun of the relative clause. This is a crosslinguistically rare but plausible development. The different languages studied in this paper show variation in the properties indicative of the status of the ‘*moto* construction’, which reflects the different stages of grammaticalization. Finally, we show how contact-induced grammaticalization is a likely factor in the development of *moto* as a focus marker.

Keywords: focus, cleft, Lingala, Kikongo, Bantu, grammaticalization

DOI 10.1515/ling-2014-0033

1 Introduction

In the vehicular Bantu language Kituba¹ we find sentences like the one in (1), where the word *muntu* ‘person’ can appear after the subject of a canonical SVO sentence. The same construction with *moto* ‘person’ and the same focus

¹ Kituba is the vehicular language used in the south-west of the Democratic Republic of Congo and in the south of Congo-Brazzaville. It is a Kikongo-based creole, and alternative names are Vehicular Kikongo, Monokutuba and Kikongo ya leta.

***Corresponding author: Jenneke van der Wal:** Faculty of Modern and Medieval Languages, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DA, United Kingdom.

E-mail: jennekevanderwal@gmail.com

Jacky Maniacky: Royal Museum for Central Africa, Leuvensesteenweg 13, 3080 Tervuren, Belgium. E-mail: jacky.maniacky@africamuseum.be

interpretation is also found in the neighboring language Lingala, as shown in (2).

(1) Kituba²

pusu muntu me kudya mbisi
 cat person PRF eat fish
 ‘It’s the cat that has eaten the fish.’

(2) Lingala

nyáu moto azalí kolía mbísi
 9.cat 1.person 1.be.PRF 15.eat 9.fish
 ‘It’s the cat that is eating the fish.’
 (C36d)³

As indicated by the translations, the element preceding ‘person’ is highlighted in these sentences. For Lingala and Kikongo we argue that this construction with ‘person’ is a focus construction, where ‘person’ functions as a focus marker. The details of the meaning and the contexts for use are discussed in more detail below to support our analysis as a focus construction.

Although one may be inclined to think of the construction as a result of creolization processes, being used in Kituba and Lingala, in fact the construction surfaces in many original languages in the region. We have encountered similar constructions in Lomongo (C60), Embuun (B87), Embosi (C25), Embudza (C37), Lingombe (C41), Otetela (C71), Kimanyaanga (H16b), Kiyoombe (H16c), Bomboma (C411, Toronzoni 2004), Sengele (C33, Niyibizi 1987), Iwonk (C85), and Tswa (C611), some of which are illustrated below.⁴

2 Abbreviations and symbols used in this paper: * (ungrammatical), # (inappropriate), ANT (anterior), APL (animate plural), ASG (animate singular), CAUS (causative), CONN (connective), COP (copula), DEM (demonstrative), FOC (focus), FUT (future tense), IMPF (imperfective), ISG (inanimate singular), LOC (locative), NEG (negative), OREL (object relative), PAST (past tense), PL (plural), POSS (possessive), PRES (present tense), PRF (perfective), PRO (pronoun), PROG (progressive), SG (singular). Numbers refer to Bantu noun classes, and to persons when followed by SG/PL; high tones are indicated by an acute accent.

3 For each language we give the classification number according to Maho’s (2009) updated classification of Guthrie (1948).

4 When no source is mentioned, the data come from elicitation sessions with our informants, either in Brussels or via telephone, usually one informant for each of the languages, but more for Kituba and Lingala (see acknowledgements). We have not controlled for the influence of the speakers being multilingual, though all of them were aware that we were asking for a specific language, and some could compare between the languages they speak.

(3) a. Iwonk

wé **moot** íunkudzí mayoŋ mámi
 2SG MOTO 1.eat.PRES.NARR 6.maize 6.POSS.1SG

‘It’s you who eats my maize.’

(C85, Tete Wer Sey 1975: 87)

b. Mbuun

taar **muur** ateema baan
 1.father MOTO 1.call.PRES 2.children

‘It’s father who is calling the children.’

(B87)

c. Kimanyaanga

ngoombe **muuntu** wetibooka
 9.cow MOTO 1.cry.PRG

‘It’s the cow that cries.’

(H16b)

In the rest of the paper, we speak of this type of sentences with ‘person’ as the *moto construction* (MC), referring to the Lingala word for ‘person’. We use the word *moto* to represent words meaning ‘person’ from other languages as well (like *muuntu*, *muur*, *bonto* etc.). We gloss the grammatical word for ‘person’ as MOTO, because on the one hand it is obvious that it does not refer to a person in the MC and on the other hand, the function of the MC is sometimes still unclear and not identical crosslinguistically, so a gloss as FOC is not warranted.

We start the paper by discussing the synchronic properties of the MC in Kituba and Lingala (Section 2). With this background, we explore the origin of the MC in Section 3, starting with a general discussion on the grammaticalization processes involved, and then testing the various languages in our sample for the parameters that are indicative of the grammaticalization path from biclausal cleft to monoclausal focus construction. Section 4 discusses the subject/object asymmetry that is present in the MC in some of the languages. In Section 5 we further discuss the historical development of the MC in the various related languages, proposing that this is a case of contact-induced grammaticalization. Finally, Section 6 draws conclusions and indicates areas for further research.

2 Synchronic properties in Kituba and Lingala

In order to describe the MC as a focus construction in Kituba and Lingala, we show first that *moto* is no longer a lexical content word when used in the MC, and second that the element preceding *moto* is indeed focused.

2.1 Moto in lexical and grammatical functions

The word ‘person’ as used in the MC has the same form as the lexical word for ‘person’, as shown in (4) and (6). It also has the plural form, just like the plural of the lexical noun, as shown (5) and (7).

(4) Kituba

a. *mfumu ya bawu ke muntu ya mbote*
 chief CONN 3PL.PRO COP person CONN good(ness)

‘Their chief is a good person.’

b. *mwana muntu me bula mutopi*
 child MOTO PRF hit football

‘It’s the child who have played football.’

(5) a. *bantu zole me kwiza kupesa nge mbongo*
 person two PRF come give.INF 2SG money

‘Two people have come to give you money.’

b. *bana bantu me yobila*
 2.children MOTO.PL PRF bathe

‘It’s the children who have washed themselves.’

(6) Lingala

a. *namóní moto mǎkó*
 1SG.see.PRF person one

‘I have seen one person / I have seen someone.’

b. *mwána moto azalí kobénga mbwá*
 1.child MOTO ASG.be.PRF 15.call 9.dog

‘It’s the child who is calling the dog.’

(7) a. *bato bazaláki ebelé na ndáko ya nzámbe*
 2.people APL.be.PAST many LOC 9.house 9.CONN 1.God

‘There were many people in the church.’

b. *bána bato bazalí kobénga mbwá*
 2.children MOTO.PL APL.be.PRF 15.call 9.dog

‘It’s the children who are calling the dog.’

It is clear that in the MC, *moto* is used as a grammatical marker rather than a noun. First, it does not function as an argument of the verb. Second, the grammatical *moto* can co-occur with the lexical *moto*, as shown in (8) and (9).

(8) Kituba

muntu muntu me yimbila
 person MOTO PRF sing
 ‘It’s a person who has sung (not a bird).’

(9) Lingala

moto moto aléí mbísi, nyáu té
 1.person MOTO ASG.eat.PRF fish 9.cat NEG
 ‘It’s a person/human being who has eaten the fish, not a cat.’

Third, the grammatical *moto* is no longer referential and as such, it is insensitive to the semantic features of person and animacy. The element preceding *moto* can be human, which matches with the original features of *moto* as a referential noun. This was already shown in (4) to (7) above, where the subject noun is ‘children’. However, not only humans are allowed: animals can also appear in the MC (see (1) and (2) above), and even inanimate referents can occur before *moto*, as illustrated in (10) and (11). In general, there is still a number distinction, but Meeuwis (1998: 43) reports the use of plural referents with the singular or invariant form *moto* (not *bato*) in Lingala (12).

(10) Kituba

inti muntu me kubwa
 tree MOTO PRF fall.INF
 ‘It’s the tree that has fallen.’

(11) Lingala

nzeté moto ekwéí kúná
 tree moto ISG.fall.PRF there
 ‘It’s a tree that has fallen down there.’

(12) Lingala

bamibáli moto/bato basómbí ndáko
 2.man MOTO/MOTO.PL APL.buy.PRF 9.house
 ‘It’s the men who have bought the house.’
 (Meeuwis 1998: 43)

This demonstrates that *moto* functions as a grammatical element in the MC. Although it has not changed formally, that is, it has not undergone phonological attrition and is still an independent word, it is decategorized from a noun to a grammatical marker, and it is desemanticized, no longer being referential or sensitive to animacy. These are important characteristics of grammaticalization

(Heine and Kuteva 2002), defined as the development from a lexical item – in this case the noun *moto* ‘person’ – to a grammatical item – the focus marker *moto* (as we argue below). However, a lexical item never simply undergoes grammaticalization by itself. It is always the item in a certain context or construction that develops grammatical functions. Therefore, we shall consider the grammaticalization process of the whole MC in Section 3, after having argued that Kituba and Lingala have indeed undergone (part of) the grammaticalization process towards a focus construction.

2.2 The MC as a focus construction

Apart from the translation of the MC as a cleft in English or French, there are four other arguments that show the focused interpretation of the referent preceding *moto*. First, *moto* is obligatory⁵ in the answer to a subject question, as in (13b). If *moto* were not present in this example, the interpretation would be one where the verb phrase constitutes the new information, as in (13c).

(13) Lingala

- a. *náni azalí kobénga bána?*
 who ASG.be.PRF 15.call 2.children
 ‘Who is calling the children?’
- b. *tatá moto azalí kobénga bána*
 father MOTO ASG.be.PRF 15.call 2.children
 ‘FATHER is calling the children.’
- c. # *tatá azalí kobénga bána*
 father ASG.be.PRF 15.call 2.children
 ‘Father is calling the children.’ (answer to ‘what is father doing?’)

Second, and related to the interpretation and context of (13c), the MC cannot be used with topics. If the referent before *moto* is already mentioned before, and is active in the mind of the addressee, it should not occur in the MC, as shown in (14a). The same applies to contrasted topics, which shows that the MC does not express a more general feature like contrast, either. This is exemplified in (14b), where *tata* ‘father’ and *mama* ‘mother’ are contrasted topics, and the use of *moto* is ungrammatical.

⁵ Except if another focus strategy can be used (see the brief mention of *mbe* and *nde* in Section 6).

(14) Lingala

a. context: what is father doing? what about father?

tatá moto azalí kobénga bána
 father MOTO ASG.be.PRF 15.call 2.children
 ‘It is father who is calling the children.’

b. *tatá, (*moto) azalí kobénga bána; mamá, (yé)*
 father (MOTO) ASG.be.PRF 15.call 2.children; mother (her)
*(*moto) azalí kobénga banyama*
 (MOTO) ASG.be.PRF 15.call 2.animal

‘Father is calling the children; mother (as for her) she is calling the animals.’

Third, referents modified by the focus particle ‘only’ are allowed to occur in the MC, as shown in (15) and (17). Notably, when modified with the (scalar) focus particle ‘even’ the construction is ungrammatical, as shown in (16) and (18). This suggests an exclusive reading of the focused element: the predicate is only true for the mentioned referent and not for alternatives (cf. Rooth’s [1992, 1996] theory of focus as evoking alternatives; see among others Krifka [2007]; Szabolcsi [1981]; É. Kiss [1998; 2006, 2010]; and Kenesei [2006] on the identificational interpretation of the preverbal position in Hungarian, and van der Wal [2011] for a discussion of exclusive focus in the Bantu language Makhuwa). This is also the interpretation described by Hulstaert (1966: 574) for the MC in Lomongo: “L’addition de *bonto* comporte une mise en relief: il s’agit bien de cet être déterminé à l’exclusion de tout autre. Sans *bonto*, le sens reste général.” [Adding *bonto* [Lomongo equivalent of MOTO] puts an accentuation: it is really about this particular being to the exclusion of all others. Without *bonto*, the interpretation remains general.]

(15) Kituba

kaka mwana muntu me dila
 only 1.child MOTO PRF cry
 ‘(It’s) only the child (who) has cried’

(16) Lingala

**ata mbuta muntu me dila*
 even old.person MOTO PRF cry
 int. ‘(It’s) even the old man (who) has cried.’

(17) Lingala

káka moto molaí moto akokí kosála óyo
 only 1.person 1.tall MOTO ASG.can.PRF 15.do this
 ‘(It’s) only a tall person (who) can do this.’

- (18) *atá moto molái (*moto) akokí kosála óyo*
 even 1.person 1.tall moto ASG.can.PRF 15.do this
 ‘(It’s) even a tall person (who) can do this.’

An exclusive focus reading is also found to be the difference between subject questions with and without *moto*. For some speakers of Lingala and Kituba the presence of *moto* in a question like (19) or (20) triggers an identificational reading of the focused referent, implying that the answer is to be taken from a restricted set. The question without *moto* just asks for any answer, without any presupposition, whereas the same question with *moto* implies that the speaker already has a certain set of people in mind and just wants to know who of them did it. The MC is also used in alternative questions like (21), where one of the overtly given alternatives must be chosen.

- (19) Kituba
nani (muntu) me kudya dimpa?
 who MOTO PRF eat bread
 ‘Who has eaten the bread?’

- (20) Lingala
náni (moto) apasólí kópo?
 who MOTO ASG.eat.PRF glass
 ‘Who broke the glass?’

- (21) Lingala
nyáu tó mbólókó moto ayíbí biléi?
 9.cat or 9.antilope.sp MOTO ASG.steal.PRF 8.food
 ‘Is it the cat or the antilope that has stolen food?’

These uses show the focused interpretation of the referent preceding *moto*. That is, if the MC is used, the referent before *moto* is in focus, but the inverse does not hold: the MC is not obligatorily used when expressing focus. One factor is that there seems to be a subject-object asymmetry for the MC, as discussed in Section 4, and another factor is the existence of an alternative focus strategy with the focus marker *nde* or *mbe*, as briefly addressed in Section 6. We conclude for now that there exists a strategy in Kituba and similarly in Lingala whereby the referent in initial position is followed by the grammaticalized word *moto*, resulting in a focused interpretation.

3 Development from cleft construction

This section discusses the question how such a focus construction came about. Our hypothesis is that the MC originated in a cleft construction ('it is X, the person who did Y'), which is corroborated by the general linguistic literature on the diachronic development of focus constructions, and the synchronic variation in the properties of the MC in neighboring languages.

3.1 Hypothesis and theoretical background

We propose that the origin of the MC in Kituba and Lingala is a cleft construction. A transparent cleft construction consists of two clauses: one containing a nominal predicate and one containing a free relative clause. The abstract representation and the example in (22) schematically show the biclausal cleft and the development into a focus construction (without showing the intermediate stages between the two).⁶

- (22) [copula – NP] – [relative clause] > [NPfocus Verb]
 it is Maud – (the person/one) who made pancakes > MAUD made pancakes

Applied to the MC, the schema looks as in (23): there is a noun which is made predicative (by a copula, or a tonal pattern), and this predicative clause is followed by a relative clause headed by the noun *moto*. One of the indications for this origin is the fact that the MC can optionally contain a copula before the focused noun, like *ezali* in (24). At some point, the construction must have been reanalyzed, resulting in a different constituency (as shown in the second part of (23)): *moto* is no longer the head of the relative clause, but pertains to the noun phrase as a nominal modifier, or is a Focus head in the left periphery of the sentence. Note that the surface structure (linear word order) is the same in the cleft and the focus construction: [copula] [focused noun] [moto] [verb phrase], which is a necessary precondition for reanalysis (Hopper and Traugott 2003).⁷

⁶ A cleft is a crosslinguistically very well known strategy to express focus. As such, it could also be called a focus construction. In this paper, we use the term 'focus construction' to refer to the monoclausal MC which expresses focus.

⁷ In the rest of the paper, we refer to the first noun in such a construction as the focused element or focused noun, and the element developing as focus marker we refer to as the head noun.

(23) [copula NP] [moto V-relative] > [[(copula) NP moto] V]

(24) Lingala

(ezalí) nyáú moto azalí kolía mbísi
 COP 9.cat MOTO ASG.be.PRF 15.eat 9.fish
 ‘It’s the cat that is eating the fish.’

The relative clause in such a cleft construction is free and non-restrictive, that is, it forms a nominal phrase by itself (‘the person/one who made the pancakes’) and it is simply adjoined to the predicative noun. Such a relative clause is headed by a definite nominal element, a DP, which gives it a maximality presupposition. The referent of the relative clause is thus presented as the complete set of referents to which the predicate applies (‘who eats the fish’ = ‘all who eat the fish’). In a cleft, this maximal group of referents is equated to the referent in the predicate, which renders an identificational and exclusive reading. For the example with the cat and the fish, the interpretation can be represented as ‘there is a maximal individual that eats fish and that individual is identical to the cat’. In the transparent biclausal cleft, this exclusive focus interpretation is brought about by the combination of a free relative clause and a nominal predicate. It stands to reason that the MC developing from it retains the same interpretation as exclusive or exhaustive focus, at least in the first stages after reanalysis.

Knowing that the original cleft already expressed focus on the first (predicative) noun, we can better understand the reanalysis of *moto*. In the biclausal cleft, the focused interpretation is the result of the whole construction of predication and free relative clause. At some point this focus interpretation must have been linked to the word *moto* (in combination with a sentence-initial position) rather than to the whole sentence construction. This is a process that Croft (2000) calls “hypoanalysis.”

In hypoanalysis, the listener reanalyzes a contextual semantic/functional property as an inherent property of the syntactic unit. In the reanalysis, the inherent property of the context [...] is then attributed to the syntactic unit, and so the syntactic unit in question gains a new meaning or function. (Croft 2000: 126)

Hypoanalysis is more likely to happen to pieces of morphology that do not seem to have a function anymore (cf. Willis [2010] “obsolescent morphology”), a process also known as exaptation (Vincent 1995) or functional renewal (Brinton and Stein 1995). In the MC, it is likely that the head noun of the relative clause (*moto*) did not have a clear function anymore, especially if several of the monoclausal syntactic properties were already present. This is the case, for example, if the copula became optional and if the verb was not recognizable as relative –

or rather, if the difference between relative and non-relative verb forms is not overtly marked. What remains seems to be a canonical, monoclausal SVO sentence with a word *moto* intervening between subject and verb. The focus meaning that was once the result of a productive combination of a free relative clause and identification by predication is now connected to the presence of the word *moto*.

In the grammaticalization from biclausal cleft to monoclausal focus construction it is usually the copula or a relativizer (like the complementizer *that* or *which*) that develops into a focus marker. Heine and Reh (1984) sketch a general scenario for the development from cleft to strongly grammaticalized term focus, based on several African languages. They propose four stages, where in the first stage there is a cleft structure like [NP/PP – copula – subordinate clause]. In stage II “the copula is desemanticized to a focus marker” (p. 181). Harris and Campbell (1995) also note some universal changes in the path cleft > focus construction, one of which is that “a discourse marker [...] is formed from some combination of (i) the copula, (ii) the relativizer, and (iii) the expletive pronoun” (Harris and Campbell 1995: 167). If a focus construction is derived from a cleft, they say, this is one of the changes that will occur. Both Heine and Reh (1984) and Harris and Campbell (1995) refer to Somali – among many other exemplified languages – to illustrate the origin of the focus marker *baa* as a copula (25). We mention it here just to give an example of the reanalysis of the copula as a focus marker (though see Tosco [2002] for arguments that *baa* is not a pure focus marker).

(25) Somali

- a. *adiga baa muuska cunaya*
 2SG.PRO FOC banana eat
 ‘YOU are eating the banana.’

(Antinucci and Publielli 1984: 19, quoted in Harris and Campbell 1995: 159)

- b. *ma Cali baa*
 Q Ali COP
 ‘Is it Ali?’

(Heine and Reh 1984: 171, referring to Lamberti 1982)

Remarkably, the focus marker in the MC is not the original copula. Instead, it is the internal head noun of the relative clause that is desemanticized and reanalyzed as a focus marker. To our knowledge, this has not been reported so far, but should actually not be too surprising, given that it is one of the linguistic elements present in a cleft (the others being the expletive, copula and relative

marker). If an internal head noun is present in the relative clause of the cleft, it too is a good candidate to become reanalyzed and be recruited for the new function of focus marker.

One of the questions is whether the development from head noun to focus marker requires intermediate steps. As possible paths of grammaticalization from the word 'person' in African languages, Heine et al. (1993) do not mention the development to focus marker, but their list includes a development to an impersonal marker ('one'). If the head noun is taken to be an impersonal pronoun, this could be an intermediate step between the lexical use of the referential noun 'person' and the grammaticalized use as a focus marker.

A second intermediate step would be from impersonal pronoun 'one' to relative clause marker. Although this, too, is a possibility, we believe that it is not necessary to assume such an intermediate step, and in fact, we do not have clear evidence in favor of it. In none of the languages can *moto* be used in a simple relative clause, and hence it cannot be analyzed as a general relative marker. If *moto* first develops into a marker of the relative clause that forms part of the original cleft, then it was/is restricted to that construction, and we cannot be sure of the syntactic status of *moto* as a head noun or a reanalyzed relative marker. We shall refer to this relative marker analysis in the following sections, but we note here that it is not necessary to assume the intermediate stages.

As already brought up in this discussion, the reanalysis of a cleft to a focus construction involves the change from a biclausal to a monoclausal structure. The biclausal cleft consists of a main clause containing a nominal predicate (usually with a copula) and a subordinate clause with the structure of a relative clause, whereas the monoclausal focus construction is one clause with one verb and a focused constituent. Harris and Campbell (1995) list three stages in the development from biclausal cleft to monoclausal focus construction (see also Harris 2001).

(26) Development cleft > focus construction (Harris and Campbell 1995: 166)

Stage 1: The structure has all of the superficial characteristics of a biclausal structure and none of the characteristics of a monoclausal one.

Stage 2: The structure gradually acquires some characteristics of a monoclausal structure and retains some characteristics of a biclausal one.

Stage 3: The structure has all of the characteristics of a monoclausal structure and no characteristics of a biclausal one.

In order to determine in which stage a certain construction in a certain language is, we should first know which properties are taken to be characteristic of a monoclausal structure or a biclausal structure, and these properties should be exam-

ined in the specific construction in the language. Harris and Campbell identify the changes in (27) that occur in Stage 2, to which Jendraschek (2009) adds the monoclausal properties in (28).

- (27) Changes biclausal > monoclausal (Harris and Campbell 1995: 166, 167)
- changing the case of the focused constituent
 - changing the form of the focus marker to look less like the copula or relativizer
 - dropping the copula or relativizer altogether
 - ceasing to use a special verb form
 - (re)introducing agreement according to monoclausal structure
 - reordering of constituents
- (28) Additional monoclausal properties (Jendraschek 2009)
- one focus marking is used for different scopes (focus projection)
 - a focused constituent can break up the original relative clause
 - there is a unified prosodic contour

Not all of these properties are relevant or even applicable to the Bantu languages in this study. Hence, we have selected the criteria we found useful in researching the stages of grammaticalization of the various Bantu languages, which we present and explain in (29). Although prosody could be a further useful property, we did not include it in this research, because of the lack of sufficient prosodic data for the languages described.

- (29) Properties in the development from biclausal to monoclausal status in the *moto* construction

1. Semantic properties of the focused element

In a transparent cleft, *moto* can only be used if the focused element is human and singular. The more developed the construction is, the less constrained is the choice of the focused noun in terms of semantic features, so that inanimates can also occur in the MC.

2. Form of the head noun

If a cleft is formed with a non-human referent, the head noun of the relative clause, which refers to that referent, should correspond in semantic features like plurality and animacy in the initial stages. That is, not only the light semantic noun ‘person’ participates in the MC, but also ‘thing’ or ‘animal’. The more desemantized the head noun becomes and the more developed the construction, the less diverse the head noun is. It becomes specialized to *moto*.

3. Presence of a copula (or other marker of nominal predication)
In the least developed case (i.e., a biclausal cleft) an overt copula or a tonal pattern expressing predication is always present, in more grammaticalized cases it is optional and eventually the copula is never used.
4. Agreement on the verb
In most Bantu languages the verb agrees with the subject. In the initial stages (I), the verb in the relative clause of a subject cleft agrees with the head noun (HN) in the free relative clause, and not with the focused constituent (FOC), which is in another clause. In the further developed monoclausal structure (II) the verb agrees with the focused subject.
 - I. α -FOC β -HN β -verb
 - II. α -FOC β -HN α -verb
5. Marking of the relative verb
In a biclausal cleft the relative clause is somehow marked as relative, which in Bantu languages often happens by means of a relative conjugation of the verb, indicated by either an affix or a different tonal pattern. In the monoclausal structure the segmental and tonal form of the verb are no longer relative, and take the form of the non-relative conjugations.

Note that the first two properties concern semantic bleaching (Heine et al. 1991), a process associated with grammaticalization and often seen as a prerequisite for grammaticalization (but see Traugott [1988] on the independence of semantic bleaching and pragmatic strengthening). In semantic bleaching, the more concrete and specific meaning of a word – in our case, the reference to a human being – changes to a more general meaning – in our case, generalizing from human to animate to unspecified entity. Although in some way the possibilities of the (original) head noun and the focused element seem to be two sides of the same coin, we discuss them separately, because they are clearly two parameters in the overall change from cleft to focus construction and languages show a continuum in both parameters.⁸

⁸ Note that our suggestion that there is a continuum in semantic change is not incompatible with Eckardt's (2006) ideas that semantic changes are small but discrete steps. A speaker will at a given point always have certain semantic specifications for *moto*, which can be pragmatically enriched until reaching a pragmatic overload, at which point the semantics of the word and/or construction will be extended (e.g. to include animals) or changed (e.g. to a general notion of 'entity').

In the following subsections, we discuss all the properties in (29) for the languages Kituba, Lingala, Kimanyaanga, Kiyoombe, Lomongo and Mbuun.⁹ We first discuss properties 1 and 2 separately in Section 3.2, as they concern the semantic change rather than the syntactic status as monoclausal or biclausal construction, which is subsequently discussed in 3.3. Although all languages display the use of a construction with *moto* conveying a focused interpretation, the diversity found for the properties in (29) shows that the languages are in different stages along the cline from biclausal cleft to monoclausal focus construction. The synchronic crosslinguistic variation thus attests to the different stages in a diachronic development.

3.2 Semantic bleaching

With respect to the first property, we have not found any language where the focused element is restricted to ‘only human’, but in some languages the possible semantic features are more constrained than in others. In Kimanyaanga, inanimates and locatives are not allowed to appear in the MC (30). Although animals are not ungrammatical in the MC, it was pointed out that this seems like a personification (31). Furthermore, it is impossible to focus ‘person’ (without any modifier) in the MC, that is, we cannot have *muuntu muuntu*, as in (32). This constitutes the most restrictive system in terms of the properties of the focused noun.

(30) Kimanyaanga

a. **n'ti muuntu wabwiidi vazulwanzo*
 3.tree MOTO 1.fall.PAST 16.above.CONN.9.house
 int. ‘It’s a tree that fell down on the house.’

b. **kukati kwa mbaanza muuntu wena n'kiinzi*
 17.inside CONN.17. 9.town MOTO 1.be.PRES 3.feast
 int. ‘It’s in the center of the town that there is a party.’

(31) *mbumba muuntu wadidi mbizi*

9.cat MOTO 1.fall.PAST 9.fish
 ‘It’s the cat that ate the fish.’

⁹ In this section we limit ourselves to constructions involving a focused subject. The differences with objects are discussed in Section 4.

- (32) **muuntu muuntu ayiimbidi*
 1.person MOTO 1.sing.PAST
 int. 'It's a person who sang.' (not a bird)

The next most restrictive language, after Kimanyaanga, is Ngombe. The focused noun in Ngombe can be human or animate, but not inanimate or locative (33).

- (33) Ngombe
- a. (*édī*) *bengé bato báβomí mbwá*
 (COP) 2.child MOTO.PL 2.hit.PRF 9.dog
 'It's the children who have hit the dog.'
- b. (*édī*) *mbongó bato bápútákákí iyǎ*
 (COP) 10.elephant MOTO.PL 2.step.PAST here
 'It's the elephants who have stepped here.'
- c. (*édī*) *moto moto omowói*
 (COP) 1.person MOTO 1.1OM.kill.PRF
 'It's a person who killed him.'
- d. *(*édī*) *molé moto omowói*
 (COP) 3.tree MOTO 3.1OM.kill.PRF
 int. 'It's the tree that killed him.'
- e. *(*t*)*á ndáko moto á-díí bopélé*
 LOC house MOTO LOC?.be.PRF beautiful
 int. 'It's in the house that it's beautiful' /
 'It's the inside of the house that is beautiful.'

In Kiyoombe and Lingala, the MC is less restricted, allowing humans, animals and inanimates to occur with *moto*, but excluding locatives; see (39a) and (40a) further below, and (11) above.

Finally, in Mbuun and Kituba no restrictions have been found as to the animacy of the focused element, as shown in (34) and (35). Even the locative context is accepted. Table 1 gives an overview of the properties for each language.

Table 1: Animacy restrictions on the focused noun

Property	Kimanyaanga	Ngombe	Kiyoombe	Lingala	Mongo	Mbuun	Kituba
Human	+/-	+	+	+	+	+	+
Animate	+/-	+/-	+	+	+	+	+
inanimate	-	-	+	+	+	+	+
Locative	-	-	-	-	+?	+	+

(34) Mbuun

- a. *mwan muur awim*
 1.child MOTO 1.sing.PRES
 ‘It’s a child who sings.’
- b. *okam muur ade mbits a mats*
 3.cat MOTO 1.eat.PRF 9.meat CONN 6.water
 ‘It is the cat who ate the fish.’
- c. *eba muur abwi*
 7.palm.tree MOTO 1.fall.PRF
 ‘It’s a palm tree that has fallen.’
- d. *ibuu la bul muur apa bis mweey*
 5centre CONN.5 14.town MOTO 1.give.PRES 1PL.PRO 3.fear
 ‘It’s the centre of town that scares us.’

(35) Kituba

- a. *mwana muntu me bula mutopi*
 child MOTO PRF hit football
 ‘It’s the child who has played football.’
- b. *pusu muntu me kudya mbisi*
 cat MOTO PRF eat fish
 ‘It’s the cat that has eaten fish.’
- c. *inti muntu me kubwa*
 tree MOTO PRF fall
 ‘It’s the tree that falls.’
- d. *Matonge muntu ke kinisaka munu*
 Matonge MOTO IPFV make.dance me
 ‘It’s Matonge that makes me dance.’¹⁰

It must be noted that this table makes the facts appear more clearly and discrete than the (socio)linguistic reality. The acceptability of the non-human nouns in the MC differs per region and even per speaker within one language. We have the impression that all speakers of Lingala find the MC acceptable with a human focused noun, many would also use it with a “likeable” animal that is close to people (like cats and dogs, but not rats or snakes), and some accept the MC with all animals and inanimates whereas others strongly dislike the use of inanimates in the MC. Hence, it seems that the reanalysis of *moto* has occurred in a part of the speakers and that actualization is taking place, with the more grammaticalized

¹⁰ Matonge is a neighbourhood in Kinshasa (and in Brussels), hence locative.

MC spreading in the language. This attests to the gradualness of grammaticalization, see e.g., Lichtenberk (1991), Croft (2000).

With respect to the second parameter, the form of the head noun, there is much variation as well. In Kiyoombe and Kimanyaanga not just *moto*, but different head nouns are available in the MC. In Kimanyaanga, if the focused noun is an animal, the head noun is *buulu* ‘animal’ (36), and if the focused noun is inanimate, the head noun is *kiima* ‘thing’ (37).

(36) Kimanyaanga

- a. *ngoombe bulu kyetidya matiiti*
 9.cow 7.animal 7.eat.PROG 6.grass
 ‘It’s the cow that eats the grass.’
- b. **ngoombe muuntu yetidya matiiti*
 9.cow MOTO 9.eat.PROG 6.grass
- c. **ngoombe kiima kyetidya matiiti*
 9.cow 7.thing 7.eat.PROG 6.grass

(37) *n’ti kiima kyabwiidi vazulwanzo*
 3.tree thing 7.fall.PAST 16.above.CONN.9.house
 ‘It’s a tree that fell on the house.’(38) *baana baantu etibula nkweezo*
 2.children MOTO.PL 2.beat.PROG 9.football
 ‘It’s the child who plays football.’

In Kiyoombe, there seems to be a choice between *moto* and ‘thing’ for non-human focused elements, as shown in (39) and (40).

(39) Kiyoombe

- a. *wwááyí múútú voondidi thúúmbi*
 7.cat MOTO 1-kill.PRF 9.mouse
- b. *wwááyí ‘bbuulu kivoondidi thúúmbi*
 7.cat 7.animal 7-kill.PRF 9.mouse
- c. *wwááyí kiima kivoondidi thúúmbi*
 7.cat 7.thing 7-kill.PRF 9.mouse
 ‘It is a cat that killed the mouse.’

(40) a. *nttí múútú voondidi báana*
 3.tree MOTO 1.kill.prf 2.children
 b. *nttí kiima kivoondidi báana*
 3.tree 7.thing 7.kill.PRF 2.children
 ‘It is a tree that killed the children.’

In both these languages the plural form of the head noun is used if the focused element is plural (41).

(41) Kiyoombe

bakyéétú báátú bánílaamba pfúúfu
 2.women MOTO.PL 2.cook.PRES 10.cassava.porridge
 'It's the women who cook *fufu*.'

Apart from Kimanyaanga and Kiyoombe, none of the languages seem to use 'animal' as the head noun/focus marker. In Ngombe it is obligatory to use 'thing' as the head noun when the focused noun is inanimate, whereas in Lomongo inanimates seem to have a choice between 'person' and 'thing'.

(42) Ngombe

ngbángé eka ekwĩ
 9.harpoon 7.thing 7.fall.PRF
 'It's the harpoon that fell.'

(43) Lomongo

a. *botámbá yǒmb' íkpaki*
 3.tree 5.thing 5.fall.PAST
 'It's the tree that fell.'
 b. *botámbá boont' ókpáki*¹¹
 3.tree moto 1.fall.PAST
 'It's the tree that fell.'

Only Lingala and Kituba are more flexible in the requirement for identity in number between the focused element and the head noun. The focused noun may be plural, whereas *moto* can remain singular, as in (44) and (45), but the other way around is never found (singular focus and plural head noun).

(44) Kituba

bakiti muntu me kubwa
 chair.PL MOTO PRF fall
 'It's the chairs that have fallen.'

¹¹ There was some discussion about this sentence, because it was naturally produced by our informant, but he thought it an illogical sentence when reflecting on the combination of 'tree' and 'person'.

(45) Lingala

bamibáli moto/bato basómbí ndáko

2.man MOTO/MOTO.PL APL.buy.PRF house

‘It’s the men who have bought the house.’

(Meeuwis 1998: 43)

In conclusion, the languages under study are found to illustrate a continuum of which semantic values are allowed as focused nouns, ranging from humans > personified/likeable animals > all animals > inanimates > locations, where ‘>’ both indicates a reversed implicational relationship (if the MC is used with a certain category, it is also used with the categories to its left) and suggests a diachronic expansion of the possibilities. The desemanticization of the head noun *moto* goes hand in hand with this development, although the match is not perfect.

An important question remains with respect to the direction of change: an extended use of *moto* in contexts where the focused noun is not human can trigger the reinterpretation and bleaching of the head noun, or vice versa: the bleaching of the head noun would allow for the use of focused nouns with varying values on the animacy scale. At this point, we cannot be sure about the answer.

Another point of interest is the interaction between semantic change and syntactic change. We come back to this point after discussing in the next section the syntactic properties as given in (29, 3–6).

3.3 Syntactic properties of biclausal and monoclausal constructions

In the first stages of the development, the focused noun is still overtly marked as predicative. This does not automatically mean that a segmental copula is used, however. In Kiyoombe, nominal predication is marked by a different tonal pattern. In (46) the difference between non-predicative and predicative form of *wwaayi* ‘cat’ consists in the presence of an initial, floating low tone. The form used in the MC also uses the predicative form, as can be seen by the low tone in (46c). In our findings so far, Kiyoombe is the only language with unambiguous tonal marking of predication. In Kimanyaanga, the tonal pattern of the focused noun in the MC appears to differ from both the subject pattern and the predication pattern (47).

(46) Kiyoombe

- a. *wwááyí* 'a cat'
 b. *`wwááyí* 'it is a cat'
 c. *`wwááyí múútú voondidi thúúmbi*
 7.cat moto 1.kill.PRF 9.mouse
 'It is a cat that killed the mouse.'

(47) Kimanyaanga

- a. *mwaana* subject of a sentence
 b. *mwáana* predication (has to be followed by a comment)
 c. *mwáána* elsewhere

(48) *mwáána múuntu wéna mu máza*

- 1.child moto REL1-be-PRES 18-LOC 6-water
 'It is the child who is in the water.' [nothing else]

In later stages, and in most of our languages, the copula is optional, as in Lomongo with *ele* (49). In none of the languages is the copula always or obligatorily absent in the MC.¹² It is important to point out that (at least in Lingala and Kituba) the copula is obligatory in nominal predication, which means that omitting it in the MC is a significant clue to its ongoing grammaticalization.

(49) Lomongo

- (*ele*) *bána banto básaná la ndembó*
 9.COP 2.children MOTO.PL 2.play with rubber/ball
 'It's the children who play football.'

(50) Kituba

- a. Joseph *(kele) longi
 Joseph COP teacher
 'Joseph is the teacher.'
 b. longi kele Joseph
 teacher COP Joseph
 'the teacher is Joseph'

12 Mbuun, Kimanyaanga and Ipala seem also to have predication contrasting tonally with some or all other contexts. In MC constructions, we see both situations: predication marking the first element (the focused element) in Mbuun like in Kiyoombe and the second (*moto*) in Kimanyaanga and Ipala. Since restrictions in the use of *moto* do not follow this distribution, the location of the predication might not play a role in the evolution of MC.

(51) Lingala

- a. *Joseph *(azalí) molakisi*
 Joseph ASG.be.PRF teacher
 ‘Joseph is the teacher.’
- b. *molakisi, ezalí Joseph*
 teacher ISG.be.PRF Joseph
 ‘the teacher (it) is Joseph’

The choices and restrictions in the form of the head noun (property 2, addressed in 3.1) are related to the fourth property: agreement on the verb. If there is a productive use of light nouns other than just ‘person’, the agreement on the verb is always with the head noun, which is the case in almost all languages. If moto is generalised further, the verb agrees with the subject, that is, with the focused noun.

The earlier stage is well illustrated in Kiyoombe: in (39a) above, the verb (*voondidi*) agrees with *muutu*, which is in class 1, even though the focused element *wwaayi* ‘cat’ is in class 7. Similarly, in (40a) the verb agrees with the head noun *múútú* in class 1 and not with the focused element *ńttí* ‘tree’ in class 3.

The later stage is found in Kituba and Lingala, where the verb always agrees with the focused element. This property could be seen as related to the fact that Kituba and Lingala, being creole languages or lingua francas, have a reduced noun class system and hence reduced agreement. Nevertheless, there is still a difference between human and non-human agreement, as shown in (52). The auxiliary *ezalí* agrees with the non-human subject (*mokobo* ‘paint’) and not with the human original head noun moto. Moreover, the language Mbudza, which is not a creole or a lingua franca, also shows agreement with the focused element (53).

(52) Lingala

- a. *mokóbo moto ezalí kokómisa bána bulé*
 paint MOTO ISG.be.PRF 15.become.CAUS 2.children blue
 ‘It’s the paint that makes the children blue.’
- b. **mokóbo moto azalí kokómisa bána bulé*
 paint MOTO ASG.be.PRF 15.become.CAUS 2.children blue

(53) Mbudza

- a. *mbai moto nkiaki lohuli*
 1SG.PRO MOTO 1SG.do.PRF 11.deal
 ‘It’s me who has made the deal.’

- b. *ima moto aholi esinza ena*
 mum MOTO 1.buy.PRF 7.dress 7.DEM
 ‘It’s mum who has bought this dress.’
 (André Epanga, personal communication)

The fifth property is the marking of the second part as relative or not. In none of the languages is the relative clause in the MC marked by a separate dedicated complementizer or particle, but the verb can be marked as relative. In the earlier stages, the verb occurs in a separate relative clause and should hence be marked as such. This is the situation we find in Kimanyaanga, Kiyoombe, and Mbuun. The verb in the MC it has the same tonal pattern as in a relative clause, as illustrated in the comparison between (54b) and (54c), and (55b) and (55c), while the tonal pattern of a non-relative verb is different, as shown in (54a) and (55a).

(54) Kimanyaanga

- a. *baana baleele*
 2.children 2.sleep.PRES
 ‘The children sleep.’
- b. *báána báleele*
 2.children REL.2.sleep.PRES
 ‘The children who sleep.’
- c. *báána báantu báleele*
 2.children MOTO.PL. REL.2.sleep.PRES
 ‘It’s the children who sleep.’

(55) Kiyoombe

- a. *múútu úntteela báana*
 1.person 1.call.PRES 2.children
 ‘The person calls the children.’
- b. *múútú úntteela báana*
 1.person REL.1.call.PRES 2.children
 ‘The person who calls the children.’
- c. *táata múútú úntteela báana*
 father MOTO REL.1.call.PRES 2.children
 ‘It is father who calls the children.’

In further stages, the verb in the MC appears as in a canonical non-relative sentence, as is the case in Kituba and Lingala. The situation in Lomongo is unclear at this point. Hulstaert (1966) describes the tonal pattern of the relative as marked by a low-high rising tone on the prefix and a high tone on the final syllable of the

verb. The same pattern is found in his examples of the MC (56). Although our consultant did pronounce the restrictive relative clauses with this rising tone on the verb, he used the non-relative tonal pattern in the MC (57).

(56) Lomongo

mbúla bont' ősenjoli bokwá

9.rain MOTO 1.REL.play salt

'It's the rain that made the salt dissolve.'

(Hulstaert 1966: 567)

(57) a. *bóna bont' ősana la ndembó*

1.child MOTO 1.play with rubber/ball

'It's the child that plays football.'

b. *bón' őnkúsana la ndembo*

1.child 1.REL.play with rubber/ball

'The child who plays football.'

We argue that in the further developed languages there is no relative construction. As an alternative still involving a relative clause (at a current or preceding stage), we could argue the following. One way to interpret these examples is to say that moto functions as the relative marker here.¹³ The tonal marking of the relative has disappeared and the original head noun now marks the sentence as relative – the direction of the probably causal relation between these two developments remaining unclear. Such an intermediate stage of either an impersonal pronoun 'one' or a relativizer (as mentioned in Section 3.1) would also account for the flexible use of moto with humans, animates, inanimates and in some languages also locatives, as well as the specialization to moto, 'thing' and 'animal' no longer being used.

There are several difficulties to prove an intermediate stage as relativizer. The most telling is that moto is never seen as a relativizer outside of the MC. In restrictive relative clauses we find either tonal marking on the verb, or a different marker, such as the (grammaticalized) demonstrative *oyo* in Lingala (58), which is also used in non-restrictive or appositive relative clauses. Furthermore, the analysis as a relative marker does not find direct support in the subject-object asymmetry discussed in Section 4.

¹³ The MC is mentioned as a relative construction by Nsuka Nkutsi (1982) and Mukash (2004), but the data they provide for 'our' languages are not accepted by any of the speakers we consulted.

(58) Lingala

ezalí nyáú óyo e-lyáki nyama
 COP cat DEM ISG-eat.PRF meat

‘it is the cat who ate the meat (as opposed to another cat)’

We could also imagine that *moto* develops as a relative clause marker (perhaps a complementizer) in addition to tonal relative marking on the verb. In that case, we cannot decide on the status of *moto* as either a head noun, which may be partly desemantized, or a relative marker. Another factor is that when *moto* is present, there is always a focus reading involved, whether due to the cleft construction or a grammaticalized focus marker. In short, if there is an intermediate stage in which *moto* functions as a relative clause marker, which is theoretically not an unlikely evolution in the chain from head noun to focus marker, this would be a development that only took place in the (subject) MC, and it remains hazy because of the gradual functional shift from head noun to focus marker.

There exist further syntactic tests to establish the mono- or biclausal status of a construction, but none of these seem to derive unambiguous results. A reviewer suggested as a test “the (im)possibility of the focus construction to occur in embedded sentences: bi-clausal clefts are illicit in embedded sentences, and most notably in restrictive relative clauses, in many languages”. We disagree that this test would give unambiguous results. This is first because the backgrounding function of a relative clause is incompatible with the focus reading of the MC, be it a biclausal or monoclausal construction; and second because clefts *can* in fact occur in subordinate clauses; see Haegeman, Meinunger and Vercauteren (2013) for arguments that the *it*-cleft should be analyzed as biclausal and is not a main clause phenomenon. We found that embedding is grammatical in both Kimanyaanga and Lingala (languages at different ends of the spectrum).

(59) Kimanyaanga

mbeenzi vo baana baantu beti yimbila
 1SG.think that 2.children moto 2SM.PROG sing

‘I think that it is the children who are singing.’

(60) *ngyeti mona woonga vo n’ti kiima kyeti bwa*

1SG.PROG see fear that tree 7.thing 7SM.PROG fall

‘I fear that it’s the tree that is falling.’

(61) Lingala

nakanísí Muriel moto amónákí mwána
 1SG.think.PRF Muriel MOTO ASG.see.PST 1.child

‘I think that MURIEL saw the child.’

Other tests mentioned by the reviewer concern binding and scope effects. The idea is that the focused element *c*-commands the rest of the sentence in a monoclausal focus construction, but not in a bi-clausal cleft.

Problematic for this test is the fact that in the biclausal cleft the coreferring head noun (i.e. *moto*) still occupies a *c*-commanding position in the relative clause, and this would still bind an anaphor or referential expression in the object position. For example, in ‘It is John (the person) who likes himself’, the head noun ‘the person’ in the second clause still binds the anaphor ‘himself’. Therefore, tests for conditions A or C will be inconclusive.

Changing the order and having the reflexive anaphor as the focused NP does not work out clearly either, e.g. in ‘It is himself that John likes’. It will be ungrammatical either because the anaphor is not bound and cannot reconstruct in the relative clause (bi-clausal) or because it is an object (monoclausal), which is also restricted in the MC (see Section 4). Even in Kituba, which freely allows objects in the MC and could hence show that the anaphor reconstructs, we find a reflexive morpheme *-di-* on the lower verb (62), which spoils the test.

- (62) *Ikele yandi mosi muntu Jean ke kudibedisa.*
 COP 3SG.PRO one MOTO Jean COP condemn.INF.REFX
 ‘It is himself that Jean condemns.’

A similarly impossible test involves possessive anaphors, where binding can potentially be seen in different quantifier scope. The idea would be that in a bi-clausal construction the focused noun would always bind a possessive pronoun in the relative clause and no ambiguity would arise. Unfortunately, this does not work for any of the following three reasons. First, universal quantifiers are excluded in an exclusive focus position, making it impossible to create a construction equivalent to *‘it is every child that likes his mother’ (cf. Van der Wal forthcoming). Second, object MCs are generally ungrammatical, making it impossible to test the equivalent to ‘it is her child that every mother likes’. Third, for the languages in which objects *can* occur in the MC, the fact that the sentence is ambiguous (63) can be explained by raising of the head noun within the relative clause, which can reconstruct and be bound by the subject in the relative clause (rather than indicating a binding relation between the possessive and the focused noun). This test would hence also be inconclusive.

- (63) Kiyoombe
bbwééndí kyándí kīima kádika muútu kázoola
 7.dog 7.POSS.1 7.thing each 1.person 1.OREL.like.PRES
 ‘It is his dog that each person loves.’ (ambiguous)

In conclusion, there is a mixture of monoclausal and biclausal properties in the languages under study, where some are further developed than others. On the “conservative” edge of the scale, the syntactic properties in Kimanyaanga and Kiyoombe point towards a biclausal structure: the focused noun is predicative, the head noun of the relative clause is variable and functions as an internal head noun (not a focus marker), and the verb agrees with the head noun within the relative clause and is still marked as relative. The MC in Kimanyaanga is thus still a biclausal cleft construction, and so is the MC in Kiyoombe, even though it seems slightly less strict in semantic terms.

Whereas Kimanyaanga and Kiyoombe can be argued to be in Harris and Campbell’s (1995) stage I, displaying only or mostly biclausal features, the other languages examined here are in stage II, because the MC behaves in some respects like a biclausal construction and in other respects like a monoclausal structure. Speculative substages of Stage 2 for *moto* could be head noun > impersonal pronoun ‘one’ > (relative marker >) focus marker. There are no systems that are fully in Stage III.

Although we have discussed the semantic changes separately from the syntactic changes, we have seen that the two do interact. Some languages show more flexibility in the animacy of the focused noun and the head noun, that is, a beginning semantic generalization, without showing all the syntactic properties associated with this process of grammaticalization. For example, Kiyoombe can use *moto* with animate focused nouns, while showing a clearly biclausal syntactic structure. This suggests that semantic generalization or bleaching indeed precedes the formal changes and syntactic reanalysis.

3.4 Other properties

Some of the monoclausal properties mentioned in (27) and (28) are not found in any of the languages, notably a) the projection of focus, b) the flexibility in position, and c) the change in the form of the focus marker. We discuss these briefly.

Ad a): The MC is in all languages used to express an exclusive focus, selecting one member of a set and excluding alternatives. It cannot be used to focus the whole sentence, the verb phrase or more than the subject or object (see also the discussion of the subject/object asymmetry below), which rejects the possibility of focus projection or focus ambiguity (cf. Selkirk 1984, 1995 on focus projection).

Ad b): As a focus marker, *moto* cannot occur in other positions in the sentence. For example, it cannot be used postverbally. The focused reading of the noun preceding *moto* may thus be a combination of the presence of the focus marker *moto* and its position in (the left periphery of) the sentence.

Ad c): The focus marker *moto* has not changed its form segmentally; in all languages it is still the same as the lexical word for ‘person’. However, in Kimanyaanga and Mbuun the tones of the lexical noun and the head noun in the MC are different. It is interesting in this respect to look at another language that uses the MC: Ipala (an unclassified Bantu language spoken in Angola). Although we do not have sufficient data to include this language into the comparison, we do know that there is a difference between the use as a lexical noun and the use in the MC. The lexical noun optionally has a pre-prefix or augment (*o-muuthu* in (64a)), depending on its function in the sentence, while in the MC it always appears without the augment (*muuthu*). The tonal marking of *muuthu* is probably the tonal pattern used for predication, which in fact suggests another analysis, as (evolving from) a copular clause: ‘the cat is the person/one who bit father’.

(64) Ipala

- a. *òmùùthù múúthù wálúpátá tààtá*
 1.person MOTO REL.1.bit father
 ‘It’s the person who bit father.’
- b. *òngáátò múúthù yálúpátá tààtá*
 9.cat MOTO REL.9.bit father
 ‘It’s the cat who bit father.’
 (J. P. Angenot, personal communication)

We conclude that in all the languages examined, the development from a cleft to a monoclausal focus construction (with *moto* as a focus marker) is underway, but has not (yet) reached the theoretical final stage of this grammaticalization process.

4 Subject/object asymmetry

So far, we have only discussed constructions in which the focused noun was the syntactic and logical subject of the sentence. However, can the MC also be used when the object is in focus? As will become clear from the discussion in this section, only Kituba seems to have no restrictions for putting an object in the MC. In the other languages objects can sometimes occur in the MC, but they are more restricted than subjects. It is not unusual to find such a subject-object asymmetry for focus marking, and it has been described for many African languages (see, for example, Fiedler et al. [2010] and references therein for West-African languages, Bokamba [1976] for Dzamba; Sabel and Zeller [2006] for Nguni; Zerbian [2006, 2007] for Northern Sotho; van der Wal [2009] for Makhuwa).

4.1 Real cleft

In the “conservative” languages, the MC is still a biclausal cleft (see Sections 3.1–3.3). The verb takes a relative form, the agreement on the verb is with the head noun of the relative clause and the head noun itself agrees in number and animacy with the focused noun. In such a transparent cleft, there is *a priori* no reason why the focused noun could not be the object of the relative clause. And indeed, we find that objects can be focused in a cleft. In Kimanyaanga, any subject or object can occur in a cleft, provided that the head noun of the relative clause agrees with it in animacy. Hence, when the focused object is human, the head noun is *múuntu* ‘person’ (65a), and when it is a non-human the head noun is *kiima* ‘thing’ as in (65b). The use of the head noun is optional, however, if a copula is present. There are two copulas, *mbe* and *yi*, which can even be combined, although either one can grammatically be used by itself. The reverse optionality also holds: if the head noun is present, the copula is optional (but note that at least one marker must be present). This variation in combinations and the resulting variation in semantico-pragmatic interpretation is a topic for further research; for now we just conclude that the focused noun can also be an object, because the construction is still a cleft.

(65) Kimanyaanga

- a. (*mbe*) (*yi*) *mfumu* (*muuntu*) *kazolele* *bula*
 (COP) (COP) 1.chief (moto) 1.want.PRF 15.hit
 ‘It’s the chief that he wants to hit.’
- b. (*mbe*) (*yi*) *nyoka* (*kiima*) *mpondele*
 (COP) (COP) 9.snake (thing) 1SG.kill.PRF
 ‘It’s a snake that I killed.’
- c. (*mbe*) (*yi*) *manga* (*kiima*) *nsuumbidi*
 (COP) (COP) 7.mango (thing) 1SG.buy.PRF
 ‘It’s a mango that I bought.’

The same holds for Ngombe (66) and Kiyoombe, as in (67) and (68), where an object can be put in a cleft, if the head noun of the relative clause agrees in animacy.

(66) Ngombe

- (*édi*) *likonja* *eka* *edukákaá* *bó*
 (COP) 5.money 7.thing 7.look.for.PROG 3PL.PRO
 ‘It’s money that they’re looking for.’

- (67) Kiyoombe
 a. *khóombó kúima phóondidi*
 9.goat 7.thing 7.kill.PRF
 ‘It’s a goat that I killed.’
 b. **khoombo muutu phoondidi*
- (68) a. *laapí kúima tsúumbidi*
 9.pen 7.thing 7.kill.PRF
 ‘It’s a pen that I bought.’
 b. **laapi muutu tsuumbidi*

4.2 Moto construction with object?

In languages where the MC is no longer a clear cleft (but *moto* has not completely developed into an independent focus marker either), we find a subject-object asymmetry. The MC is grammatical for any subject, but more restricted when the focused noun is an object. In Lingala, the MC can certainly be used for human objects (69a), but not for non-human referents, as shown in (69b).

- (69) Lingala
 a. *bána bato mamá azalí kobimisa*
 2.children MOTO.PL mother ASG.be.PRF 15.go.out.CAUS
 ‘It’s the children that mother is sending outside.’
 b. **nyama moto tatá azalí kolía*
 9.meat MOTO father ASG.be.PRF 15.eat
 int. ‘It’s meat that father is eating.’

Nevertheless, there is variation in the acceptability of the MC with objects in Lingala. If a copula is present, some speakers can still accept a construction with a focused non-human object, as in (70). The fact that the copula must be present suggests that this construction is more of a cleft than the MC in Lingala, even though the head noun does not agree in animacy or number with the focused noun (i.e., it is still *moto* and not ‘thing’).¹⁴

- (70) Lingala
 a. *ezalí nyama (moto) tatá azalí kolía*
 COP 9.meat (MOTO) father ASG.be.PRF 15.eat
 ‘It’s meat that father is eating.’

¹⁴ This could synchronically be taken as an argument to analyse *moto* as a relative marker.

- b. *ezalí bitabe (moto) tatá azalí kolía*
 COP 8.bananas (MOTO) father ASG.be.PRF 15.eat
 ‘It’s bananas that father is eating.’

Interestingly, a pseudocleft can also be formed in Lingala. Whereas in the MC Lingala does not use the semantically light words ‘thing’ and ‘animal’, in the pseudocleft these are used as the internal head noun of the relative clause. This is necessarily so, because the relative clause in a pseudocleft is referential, that is, it really refers to an animal or a thing, and the identification of this animal or thing happens in the predicate. For example, in (71c) the relative clause refers to some ‘animals’ and the predicate identifies those animals as dogs. Hence, somewhat more specific nouns may also be used in the relative clause, such as ‘food’ in (71d). This shows that the formation of a pseudocleft is independent of the grammaticalization of the MC. More importantly it confirms the well-known idea in grammaticalization theory that it is never just one word that grammaticalizes, but a word or combination of words in a certain linguistic context.¹⁵ Only in the cleft, as explained above, did the head noun *moto* develop into a focus marker, and not in the pseudocleft.

(71) Lingala

- a. *bato tatá azalí kobénga, ezalí bána*
 2.people father ASG.be.PRF 15.call COP 2.children
 ‘Who father calls, are the children’
 lit. ‘The people father calls are the children.’
- b. **bato tatá azalí kobénga, ezalí bambwá*
 2.people father ASG.be.PRF 15.call COP 2.dog
 lit. ‘The people father calls are (the) dogs.’
- c. *banyama tatá azalí kobénga, ezalí bambwá*
 2.animal father ASG.be.PRF 15.call COP 2.dog
 ‘The animals father calls are (the) dogs.’
- d. *eloko/bilia tata azalí kolía, ezalí nyama*
 2.thing/8.food father ASG.be.PRF 15.eat COP 9.meat
 ‘The thing/food father eats is meat.’

¹⁵ See for example Traugott’s (1995) definition of grammaticalisation: “grammaticalization is the process whereby lexical material *in highly constrained pragmatic and morphosyntactic contexts* becomes grammatical” (emphasis ours).

Mbuun seems to have the same pattern as Lingala: human objects can be used in the MC, but non-human objects are ungrammatical. An alternative of using ‘animal’ or ‘thing’ does not exist in this language. Notice that the fact that *moto* cannot be used in (72d) excludes an analysis of *moto* as a relative marker in (this construction in) Mbuun.

(72) Mbuun

- a. *mfum muur akwis ndzaa kobwil*
 1.chief MOTO 1.come.PRES 9.need 15.beat
 ‘It’s the chief that he wants to beat.’
- b. *mbe ikon ede*
 COP 5.banana 1SG.eat.PST
 ‘It’s a banana that I ate.’
- c. **ikon muur ede*
 5.banana MOTO 1SG.eat.PST
- d. **mbe ikon muur ede*
 COP 5.banana MOTO 1SG.eat.PST

Kituba is the only language where any element may be used in the MC, regardless of the syntactic function (subject/object) or animacy. Human and non-human objects and even locative objects can be focused in the MC, as illustrated in (73).

(73) Kituba

- a. *mfumu muntu yandi zola kubula*
 chief MOTO 3SG.PRO want beat.INF
 ‘It’s the chief (that) he wants to beat.’
- b. (*ya ke*) *bitika muntu mu(nu) me kudya*
 (COP) banana MOTO 1SG.PRO PRF eat
 ‘It’s bananas that I have eaten.’
- c. *Na Kinsasa muntu mvula me noka*
 LOC Kinshasa MOTO rain PRF rain
 ‘It is at Kinshasa that rain has fallen.’

The variation between the languages shows once more the various stages of grammaticalization, as summarized in Table 2. But why would there be a subject-object asymmetry in the languages that allow (some) objects to be used in the MC, like Lingala and Mbuun? And why is Kituba the only language that allows any syntactic function in the MC? The latter question is addressed in Section 5, and the former is the topic we turn to now.

Table 2: Overview of syntactic functions allowed in MC

	subject	object	languages
cleft	✓	✓	Kimanyaanga, Ngombe, Kiyoombe
	✓	+ human ✓ – human ✗	Lingala, Mbuun
focus construction	✓	✓	Kituba

4.3 Subjects, animacy and specialization to moto

One factor that might play a role in the subject-object asymmetry is the restriction on the syntactic function of a relativized noun. Universally speaking, it is easier to relativize subjects than objects or obliques (see the Accessibility Hierarchy of Relativization in Keenan and Comrie 1977). However, the languages discussed in this paper do have a strategy to relativize non-subjects, so it is unlikely that this was a major influence.

Another factor is the marking of subject focus. In an SVO language, most often the subject is the default topic and the object forms (part of) the new information in a canonical sentence (Lambrecht 1994). If the subject is to be interpreted as the focus, it should somehow be marked so as not to get the default reading as the topic (Lambrecht 2000, Fiedler et al. 2010). This is exactly what a cleft does: it marks the subject as the focus of the sentence. Unlike subjects, objects can be focused in their canonical (postverbal) position. Therefore, in the answer to a subject question, the cleft/MC *must* be used and moto is obligatorily present, as in Lingala (74), whereas for an answer to an object question the cleft/MC is not necessarily used (75). In fact, the object in the answer to an object question preferably occurs after the verb, as in (75c).

(74) Lingala

- a. *náni azalí kobénga bána?*
 who ASG.be.PRF 15.call 2.children
 ‘Who is calling the children?’
- b. *tatá moto azalí kobénga bána*
 father MOTO ASG.be.PRF 15.call 2.children
 ‘It’s father who is calling the children.’
- c. # *tatá azalí kobénga bána*
 father ASG.be.PRF 15.call 2.children
 ‘Father is calling the children.’

- (75) a. *tatá akobénga náni?*
 father ASG.FUT.call who
 ‘Who will father call?’
- b. (*ezalí*) *mwána moto tatá akobénga*
 (COP) 9.child MOTO father ASG.FUT.call
 ‘It’s the child that father will call.’
- c. *tatá akobénga mwána*
 father ASG.FUT.call 1.child
 ‘Father will call the child.’

If subjects cannot be focused in their canonical position, but objects can be focused in situ, it follows that the cleft as a focus marking strategy is used more often (or even predominantly) for subjects. Most likely, this unequal distribution for subjects and objects became fixed in the grammaticalization to a monoclausal focus construction (the MC). This leaves as an exception the case of Kituba, as there is no subject-object asymmetry in this language, nor is the construction used a transparent cleft. This will be discussed in Section 5.

The fact that the MC mostly contains subjects is relevant in another way: it can help us to understand better why it was *moto* ‘person’ that grammaticalized to focus marker, and not the word for ‘animal’ or ‘thing’. Typical subjects are agents, and typical agents are human. Following these hierarchies and tendencies, the focused nouns occurring in the original cleft, which are more often subjects, are more often human. The head noun of the relative clause in the cleft is coreferent with the focused noun, and hence this head noun is more frequently the human *moto* ‘person’ than ‘thing’ or ‘animal’. Frequency does have an impact on grammaticalization processes in the sense that the frequent discourse patterns become entrenched in the grammar (Ariel 2008, Bybee et al. 1994), so in the development from cleft to focus construction, the relatively high frequency of *moto* ‘person’ could form a motivation for the specialization process whereby the variation person/animal/thing becomes more and more restricted to person.

The subject agreement on the verb can be seen as another facilitating factor for the development of *moto* ‘person’ to focus marker, rather than ‘animal’ or ‘thing’. In a typical Bantu noun class system, the great majority of words for human referents is in class 1/2,¹⁶ as is the word for ‘person’. The agreement on the verb in a cleft or the MC will thus be in class 1/2 for a human referent, but crucially it cannot be determined whether the verb agrees with the head noun ‘person’ (as in a biclausal cleft) or with the focused noun (as in a monoclausal

¹⁶ Class 1 being the singular and class 2 the plural.

construction). This is illustrated in (76b), where the subject marker on the verb *bá-βomí* ‘they have hit’ is in class 2 and could thus agree with the class 2 focused noun *bengé* ‘children’ or with the class 2 head noun *bato*. This ambiguity presumably makes the transition from biclausal to monoclausal easier. For ‘thing’ and ‘animal’ the distribution of referents is different. The words/names for animals are not all in the same class as the word ‘animal’ itself, and neither are all things in the same class as the word ‘thing’, as in (77). Hence, the agreement on the verb is less often ambiguous for ‘animal’ or ‘thing’, which makes the transition to a monoclausal construction less obvious.

(76) Ngombe

- a. *bengé* *báβomí* *mbwá*
 2.children 2.hit.PRF 9.dog
 ‘The children have hit the dog.’
- b. (*édi*) *bengé* *bato* *báβomí* *mbwá*
 (COP) 2.children MOTO.PL 2.hit.PRF 9.dog
 ‘It’s the children who have hit the dog.’

- (77) a. *molé* *mómowói*
 3.tree 3.1OM.kill.PRF
 ‘The tree killed him.’
- b. (*édi*) *molé* *eka* *emowói*
 (COP) 3.tree 7.thing 7.1OM.kill.PRF
 ‘It’s the tree that killed him.’

In summary, with regard to the syntactic function of the focused noun we find three types of languages. In the first type, subjects and objects of different animacy can all be focused. Considering the relative marking on the verb and the variation in animacy of the head noun (person/animal/thing), these constructions can be considered biclausal clefts. In the other two types of languages the construction is not a clear cleft. The Lingala-type shows a subject-object asymmetry: the MC can be used to focus subjects with different values on the animacy hierarchy, but objects are only allowed if the referent is human. This restriction could (diachronically) be due to the fact that objects can also be focused in-situ (after the verb), whereas for subjects a cleft is the dedicated focus strategy. In addition, the fact that the MC is predominantly used for subjects can be linked to the specialization of the head noun/focus marker to *moto* ‘person’, via the prototypical tendencies subject=agent=human. The third type of language is Kituba, where any subject or object can be focused in the MC. The difference with the first type is that in Kituba there is no sign of the construction being a cleft.

5 Simultaneous development or spread?

Now that we have described the properties of the MC in various languages, and explained our proposal for the grammaticalization process leading from a cleft to a construction with *moto* as a focus marker, in this section we discuss the presumable historical process, on the basis of our comparative research. The main questions are why and how the grammaticalization process happened in the languages under examination.

5.1 Facilitating factors for head noun as focus marker

Although the development from cleft to focus marker is well known, the relation between the head noun and a later focus marker is not often mentioned. In this subsection we sketch the factors (if not the necessary and/or sufficient, than at least the facilitating) that contributed to the grammaticalization of the head noun, trying to explain the fact that the MC is confined to a relatively restricted area in the Bantu domain. Encountering a similar phenomenon and development in a completely different language family is not unexpected, however, and would be helpful in understanding the motivations for the grammaticalization of the head noun into a focus marker. For now, our question in this section remains: what makes the region involved more susceptible to such an evolution than other linguistic areas?

Taking back the structure copula + focused noun + head noun + relative verb, what are the necessary conditions for a cleft to start on this grammaticalization path? The key constituents being the copula, the head noun and the relative verb, we propose that the absent or relatively weak/subtle explicit marking of the copula and the relative verb facilitates the development of the head noun as a focus marker. We analyze the influences in the process for each of these.

First of all, these languages did or do not use an expletive or a relative complementizer in the original cleft construction. These two elements are a known source for focus markers developing from a cleft (Harris and Campbell 1995: 167), but this is obviously not an option if there is no complementizer in the first place, or, as Harris and Campbell (1995: 168) explain, “a language that does not use an expletive pronoun in its source cleft will not have a reflex of an expletive pronoun in its [focus marker]”.

As described above, the difference between the verb in a non-relative indicative sentence and in a relative clause is in these languages either not marked in the form at all, or marked by a different tonal pattern. This entails that the relative marking is not a strong candidate to become the focus marker. For the first case,

again stating the obvious: if there is no difference between the non-relative and the relative verb form, this ‘non-marking’ cannot be refunctionalized. For the second, although tonal marking plays an important role, it is more subtle than a segmental morpheme and more integrated into the grammatical system, being intimately linked to the verb. This means that such tonal marking is less easily reanalyzed than a segmental morpheme such as a copula or head noun, especially when it concerns an emphatic category like focus.

Concerning the copula, the same reasoning can explain for part of the languages why the copula did not become the focus marker, as in these languages the predication of the focused noun is marked by means of a different tonal pattern. A further line of thought is that in the languages of zone C the two-tone system is restricted in its use for lexical and grammatical distinctions and may therefore be less likely to also be put into use for marking focus (this also as opposed to intonation, which may always play a (additional) role). A segmental copula is synchronically also attested, but it remains largely optional in the languages involved, such as *ya ke* or *ni* in Kituba and *ezali* in Lingala.

When the marking of focus on other elements in the cleft construction (expletive, complementizer, relative verb marking, copula) is absent, optional, weakened or less likely to be reanalyzed, the head noun turns out as the optimal candidate to become reanalyzed as a focus marker. We thus propose a monodirectional implicational relation: where *moto* is (being) reanalyzed as focus marker, the other “ingredients” of the cleft construction are likely to have non-segmental marking. If this is the factor facilitating the reanalysis of the head noun, it is interesting to see whether other Bantu languages – those which do not have the MC – display the same properties as just described, and whether they use a similar cleft construction as the one from which the MC originated.

In the eastern part of the Bantu domain, different cleft strategies are found, but there is also an important difference in the behavior of the stem *-ntu* (from which *moto* is formed). The stem is well attested in Eastern Bantu languages, but compared to the western languages it has a peculiarity: the stem displays a great prefix flexibility triggering a wider semantic scope, from ‘person’ to ‘place’ via ‘thing’. This may block the adoption of one form as the focus marker, especially as prefix flexibility calls for a maintained agreement with the head noun. This is illustrated in (78), where, besides the presence of a nasal copula, *-ntu* simply takes a prefix in agreement with the head noun: class 4 and class 11, respectively.

(78) Ciluba

- a. *mítshi mmi-ntú idi amú mwí.tu*
 4.tree COP.4-ntu 4.be only 18.5.forest
 ‘It’s the trees that are only in the forest.’

- b. *lúlóní élu ndú-ntú ludi káluyi lusumíníná kúlóná*
 11.pupil DEM.11 COP.11-ntu 11.be NEG.11.do 11.effort 15.study
 ‘It’s that pupil who does not make an effort to study.’
 (L31, Nsuka Nkutsi 1982: 89)

In Western Bantu languages, where the stem *-ntu* does not have such a prefix flexibility, the focus strategies are still different from the MC (see also papers in Aboh et al. 2007 and Fiedler and Schwarz 2010). Many languages have focus strategies that synchronically do not involve a cleft, or even a segmental focus marker. One example is Kiyaka. Like in many other languages, there is no segmental difference between a noun in isolation and its predicative or focused form, but a tonal modification is applied to express what may be translated as a cleft (79d).

(79) Kiyaka

- a. *khókó* ‘(It’s) a chicken.’
 b. *khokó lǎdidi* ‘The chicken disappeared.’
 c. *khoko lǎdidi* ‘The chicken that disappeared.’
 d. *khókó lǎdidi* ‘It’s the chicken that disappeared.’
 (H31, Kidima 1990: 196)

Outside the MC area, we do find cleft constructions, but they look different from the one we propose as the origin of MC started. In a variety of Kikongo, the focus is marked with a possessive form (80). In Fang, the focused noun is followed by a copula and pronoun that mark the focus (81).

(80) Kikongo

- Francisco kwandi vondele mbw’ame ku mfinda*
 Francisco 17.POSS.3SG 3SG.kill.PAST my.dog LOC17 forest
 ‘It’s Francisco who killed my dog in the forest.’
 (H16 variant, del Fabbro and Petterlini 1977: 225)

(81) Fang

- a. *ñ-yáǵálá à-kóbá*
 1-teacher 1-talk.PRES
 ‘The teacher talks.’
 b. *ñ-yáǵálá é-nyá à-kóbá*
 1-teacher COP-1PRO 1-talk.PRES
 ‘It’s the teacher who talks.’
 (A74, Yolande Nzang Bié, personal communication)

However, we also encountered cases that are more like the MC languages. One example is Iyaa, where the tonal pattern of the focused noun is raised, as illustrated in (82). The tone change on the verb is probably a sign of a relative construction. Although the conditions appear to be the same as in our sample of MC languages (no or non-segmental marking of predication and relative verb), Iyaa does not have (the equivalent of) *moto* as a focus marker. We cannot be sure of the import this example has on our analysis, because the origin of the focus marker is unknown, and the construction may never have had an overt head noun in the first place.

(82) Iyaa

- a. *mùkò:nzì nĕmìsì mùkùlú*
 3.elbow 3.hurt.PAST 1.eldest
 ‘The elbow has hurt the eldest.’
- b. *múkó:nzì nĕmìsì mùkùlú*
 3.elbow.FOC REL?3.hurt.PAST 1.eldest
 ‘It’s the elbow that hurt the eldest.’
 (B73, Mouandza 2001: 247)

Another example is Mbosi, which uses the MC in quite a liberal way, e.g., also with inanimates and locatives. The agreement on the verb is with the focused noun (83b)–(83c), but the conjugation is not the same as in a canonical SVO sentence (83a)–(83b). This example shows the opposite situation: *moto* has developed further despite the persistence of the relative form of the verb.

(83) Mbosi

- a. *ilangi ibonzímí*
 4.bottles 4.break.PRF
 ‘The bottles broke.’
- b. *ilangi moro míibondzímí*
 4.bottles MOTO 4REL.break.PRF
 ‘It’s the bottles that broke.’
- c. *olangi moro mósóbondzímí*
 3.bottle MOTO 3REL.break.PRF
 ‘It’s the bottle that broke.’
 (C25, Guy Kouarata, personal communication)

It remains to be seen how Mbosi fits in the development from cleft to focus marker *moto*, especially with respect to the scenarios sketched in the next section (grammaticalization vs. borrowing). Although Mbosi may well be a counterexample,

we leave a more profound comparison of cleft constructions in the area to further research, and maintain our proposal that the (non-segmental) marking of the cleft construction in the MC languages is a facilitating factor for the head noun to become a focus marker.

5.2 Grammaticalization, contact and/or calquing

With these facilitating (and hence for other languages restricting) properties, we can better understand the development of the MC in this region. But does the grammaticalization process sketched in Section 3 apply to all languages that have some variant of the MC? Do these geographically related languages all follow the same grammaticalization path, or is it the result of calquing in some of the languages? In this section we consider four theoretical possibilities for the development of the MC in the various languages, presenting arguments in favor of and against them.

5.2.1 Inheritance

A first possibility is that some ancestor of the languages developed a form of MC and passed it on to the various languages. This is unlikely at the proto-language level, because such a construction is not reconstructable for proto-Bantu. It does not even involve a substantial amount of languages in the Western part of the Bantu domain. Interestingly, the MC is attested in a creole (Kituba) and in a vehicular language (Lingala), and it is a still ongoing process. Another argument against inheritance is the variety of constructions, as described in the first part of this paper. The existence of both the biclausal and the monoclausal use of the MC, with *moto* as a real focus marker in only some of those languages, rather presents evidence of a grammaticalization process.

5.2.2 Independent development

We could suppose that all languages that have a form of MC nowadays started this path of grammaticalization independently. It would appear logical that if all languages involved have roughly the same kind of cleft construction as a starting point (with optional or no copula and minimal or no relative marking), they could “automatically” enter the same grammaticalization path. This is known as “parallel development” (Crowley 1991). However, it seems too regional to assume that

the same development happened independently in each of these languages. The distribution of the languages displaying the MC suggests an areal influence.

5.2.3 Contact-induced grammaticalization

The MC as an areal feature would then be the result of a contact-induced grammaticalization. Thomason (2001: 91) notes, “Establishing the fact of contact-induced change is usually easy when the focus is on loanwords, but it can be much harder, and often impossible, with structural interference”. Nevertheless, we can imagine the following path of development: a language A starts the grammaticalization process from biclausal cleft to monoclausal focus construction, as detailed in Section 3.1. In a neighboring language B, which has a (biclausal) cleft, the grammaticalization of the cleft is triggered by contact with language A. Any other language like B, located within a synchronic buffer zone in geographical space (Kuteva 2008) is capable of undergoing the same process. This is actually what we propose for the major part of the languages we studied. The language contact is enhanced in the area by the fact that two of them, Lingala and Kituba, are vehicular, national languages.

Unfortunately, it is for now difficult to say which language in our study was language A, the starting point. There are several reasons for this. The most important one is that no historical data are available for the Bantu languages. A second reason is that the inventory of languages having the MC is not exhaustive. One has to bear in mind that the MC is not always reported in the literature on the languages investigated, even when the use of the MC is confirmed by the speakers. Therefore, languages with no mention of the MC in the literature cannot automatically be considered as not having the MC. Finally, exact origins of borrowings and interferences are usually difficult to identify when dealing with related languages.

In a contact-induced grammaticalization context, the grammaticalization process synchronically shows different levels reached in the various languages mentioned in Section 3.2. Only a few of them went further in the cline to display a monoclausal construction. Such a scenario does not exclude the possibility of straight borrowing from a relatively advanced stage in the grammaticalization cline.

5.2.4 Borrowing

The hypothesis in this scenario is that the grammaticalization process happened in one or more languages and the MC was then borrowed into other languages.

If borrowing took place, it was not just the word ‘moto’ that was borrowed, but rather the whole structure, the MC, with the corresponding word for ‘person’. Hence, it would be more adequate to talk about calquing rather than borrowing, since all languages use their own word for ‘person’ (*moto*, *muuntu*, *muur*, *bonto* ...). The calquing of the whole construction, at some point in the development, may not be the most likely scenario for most of the languages, but it is a real possibility for Kituba.

5.2.4.1 The Kituba case

In Kituba, the structure could have been borrowed and adapted after the grammaticalization process, at the point where the word for ‘person’ already has another meaning/function (i.e. focus marker). Various facts in the language support this hypothesis.

First, there is no agreement in number with the focused noun (see (44)). The fact that this agreement is maintained when the focused noun is human can be attributed to a persistence effect present in the whole (calqued) construction (Hopper 1991).

Second, there is no match in animacy with the focused noun. The use of *muntu* is general to all contexts, including locatives (35). There is no use of animal/thing. Mfoutou (2009: 63) argues for the use of the generic name agreeing in animacy with the focused noun, but this is contradicted as the ‘normal’ construction by our informants. It does show that the language has this possibility (as in any language in fact) in response to a more specific question (‘Which animal/thing ...?’), which is the transparent combination of a copular clause and an appositive relative clause.

Third, there is no other formal focus marker. This relates to the idea that Fischer (1997: 467) mentioned as one of the scenarios that borrowing may lead to: language calques, a structure to bridge a functional gap. Indeed, in our set of data, Kituba is the only language where SVO can be ambiguous between focus on the subject and a “neutral” interpretation. Besides the intonational stress on the subject, the other strategy developed is a focus marker (*muntu*), which is added to the “bare” structure to mark focus even more explicitly and to disambiguate. The reason why the equivalent of *moto* was borrowed, and not another marker (like *nde* commonly found in the Lingala area), is probably linked to the fact that the *moto*-type (biclausal sentences included) is also attested in the Kongo area, from where Kituba has mainly developed. On the contrary, *nde* as a focus marker is not attested in the Kongo area.

There is one problematic point in this argument for Kituba, which lies in the optional use of the copula. If the copula can be used, as shown in (84), then a first

question is: Where does the copula in this construction come from, if the MC is borrowed? It could be that the optional copula was still part of the MC when it was borrowed, or that it is added analogous to either a normal cleft strategy or the MC in underlying Kikongo languages. A second problem is that we actually find three indications of focus: the MC/*muuntu*, the intonation, and the copula. It would be interesting to see the precise differences in interpretation between the use of these three strategies, and their combination.

(84) Kituba

- a. *ya ke tata muntu me kwenda na kati ya inzo*
 COP father MOTO PRF go LOC inside CONN house
- b. *ya ke tata me kwenda na kati ya inzo*
 COP father PRF go LOC inside CONN house
- c. *tata muntu me kwenda na kati ya inzo*
 father MOTO PRF go LOC inside CONN house
 'It is father who went inside the house.'

As a fourth argument for the calquing hypothesis, although grammaticalization occurs in creoles like in any other language (Mufwene 2006), Kituba is more likely to have adopted this focus construction in its young history (Samarin 1991; Fehderau 1966; Mufwene 1988), than to have followed the whole grammaticalization path. Going through the grammaticalization process would imply a change in the agreement, from agreement with the head noun of the cleft to agreement of the focused noun in the monoclausal sentence. However, as a result of a creolization process, Kituba has no agreement system at all. This implies that the agreement on the verb is the same, whether agreeing with the head noun (*muntu*) or focused noun, which could actually be a facilitating factor in the adoption of a simple monoclausal focus construction.

5.2.4.2 The Lingala case

There is another language that may be a candidate in the case of a calquing of the whole MC construction: Lingala. The history of the language (Samarin 1991) shows some similarities with that of Kituba in the sense of a rapid development from local languages to a vehicular variety. In Lingala, like in Kituba, the MC can be considered as a full monoclausal focus construction. However, what is interesting in the case of Lingala compared to Kituba is that it has an agreement system even in the most simplified version (human vs. non-human subject marking on the verb). Whereas in most of the other languages studied the verb agrees with the head noun/focus marker (*moto*), in Lingala the agreement is clearly with the

focused noun (see (52)). There is no attestation of a possible agreement with *moto* in any text available. Hence, in Lingala itself there is no evidence in the verbal agreement of grammaticalization process from bi- to monoclausal construction.

5.2.4.3 Where from?

If the MC was calqued, a natural question is which language it was borrowed from. Let us first query the Kituba situation. We put forward two possible origins:

The construction could be calqued from a Kongo language. By “Kongo language” we intend a wider group of languages than the genetic classification implies: a language spoken within the area where Kituba is used as the communication language. In this area, we find other kinds of borrowings from a Kongo language to Kituba. In the cleft constructions, copulas like *mbe* (attested in Kimanyaanga) or *ni* (attested in Laadi) are sometimes used. Furthermore, the word *muntu* used in Kituba is usually morphologically identical to the word for ‘person’ in the Kongo languages.¹⁷

The other possibility is, logically, that it is calqued from another (non-Kongo) language. The main Bantu candidate is then Lingala, which is indeed the non-Kongo language from which Kituba has borrowed the highest number of items, albeit mainly in the lexical domain. As a vehicular language, Lingala is in direct contact with Kituba. Of course, the two languages do not share the same phonological system (7 vowels for Lingala, 5 for Kituba, for instance) but the observation of the borrowings from Lingala to Kituba shows that the words can be adapted, like *muntu* from *moto*. Another example is *mudinga* ‘smoke’ from Lingala *molinga*, found in Kituba as spoken at the border of the Lingala area (Swartenbroeckx 1973: 363). Finally, a borrowing from outside the Kongo languages would explain why there is no borrowing of the system with animal/thing as the head noun/focus marker, which is present in the Kongo languages.

As far as Lingala is concerned, if we assume that it calqued the monoclausal structure from another language, from the data collected among the non-vehicular languages, only Mbudza and Mbosi satisfy the criteria of agreement with the focused element, but Mbosi still has a marked relative verb form.

In this difficult task of trying to define a precise origin of the MC, we should remain open to all possibilities. In that sense, we could also presume that the directions of calquing were the other way round, that is, from Kituba into the other languages under study. However, it is very improbable that the MC as it is in Kituba developed from unspecified agreement to agreement with the head noun

¹⁷ This would actually argue for a straight borrowing and no need of calquing.

(as in various Kongo languages). Moreover, this reverse development would run counter to universal patterns of grammaticalization, which suggest a unidirectional development from biclausal to monoclausal. All the same, nowadays Lingala as well as Kituba may influence the languages spoken in their area of diffusion, and this in an even faster way than in classical language-contact.

5.3 Summary

The development from cleft to focus construction is not rare. However, the development of a focus marker from a head noun is too local to be an independent development. We suggest that this is a case of contact-induced grammaticalization, where the development started in one language and was triggered to start in other languages that were/are in contact with this language. Which language was the first to have entered the grammaticalization path is difficult – if not impossible – to determine. On the other hand, it is also not likely that all languages went through all the steps in the process. Some languages may have entered at an intermediate stage, and for Kituba specifically we argue that it has not undergone the grammaticalization from biclausal cleft to monoclausal focus construction, but that it has calqued the construction in a late stage of grammaticalization.

6 Conclusion and discussion

In this article we have shown that the word *moto* (and its variants in each language) is used in both its lexical function meaning ‘person’ and its grammatical function indicating focus. We have discussed the synchronic properties of the grammatical use in a number of related languages, concluding that some still show a biclausal cleft construction, whereas in others it is a monoclausal focus construction. Accepting that synchronic crosslinguistic variation, and possibly synchronic gradience (Traugott and Trousdale 2010), is indicative of diachronic change, our analysis exhibits new aspects in the grammaticalization path from biclausal cleft to monoclausal focus construction.

First, while we know that the copula and the relative clause marker can be reanalyzed as focus markers, this paper shows that the head noun of the relative clause can also develop into a focus marker, whether directly or indirectly via an impersonal pronoun and relative marker. It remains to be seen how many other languages in the area also display such a development, and whether it can be found in languages elsewhere in the world.

Second, grammaticalization always depends on the co-text with which an element occurs, in terms of grammatical functions and interpretations, but also on formal context. The reanalysis of the head noun was probably facilitated by the absence of segmental marking of the relative verb and the nominal predication.

Third, hypanalysis can be a gradual process. The focus function was first transparently encoded in the whole construction of a cleft, and changed to be expressed in just one word in the monoclausal construction (hypanalysis, Croft 2000). As mentioned in Section 3.1, this process is more likely to occur if the reanalyzed word does not have a clear function anymore. Although the syntactic reanalysis is abrupt (a speaker does or does not use *moto* as a focus marker (cf. Hopper and Traugott 2003: 46), the functional reanalysis is more likely to be a gradual process, where the former function disappears as the new function grows stronger. In our case, *moto* gradually gained its focus meaning while the other markers in the cleft disappear, and is eventually reanalyzed and parsed with the focused noun, and no longer as part of the relative clause. Hence, the reanalyzed word (*moto*) is never completely functionless (cf. Vincent 1995; Norde 2002).

Concerning the crosslinguistic variation in the languages examined, we suggest that it is unlikely that in such a restricted area all languages developed the same construction independently, but at the same time we do not claim that the construction was calqued after developing in one language. We propose that a process of contact-induced grammaticalization took place (Kuteva 2008), with an exception for Kituba, a relatively young creole, and possibly Lingala, where we do find reasons to say that the MC was calqued. Applying the theory of Kuteva (2008), the area involved in Central Africa behaves like a synchronic buffer zone for a grammaticalization development.

There are many loose ends in this research, and many areas for further investigation. We discuss four here. One is the link with the constructions using *-ntu* in Eastern Bantu languages. Our research started out by observing the Lingala data in the overview of relative constructions in Bantu languages by Nsuka Nkutsi (1982). He mentions the *-ntu/moto* construction as a relative strategy in Lingala, but as we have shown, it occurs with a cleft interpretation rather than as a restrictive relative clause. However, it does seem to be used in true relative clauses (not just clefts) in Bemba, which has the *-ntu* stem. This may be evidence for an intermediate stage as relative marker in the grammaticalization process.

Another question is whether there are other languages that show a similar structure and/or development. Specifically, what is the influence of the fact that lingua francas are involved in the grammaticalization process in this area?

Furthermore, we notice that some languages have an alternative focus marker (*nde* or *mbe*), as illustrated in (85). The co-existence of the MC and other focus markers raises questions on whether there is any pragmatic difference between the use of *moto* or one of the other focus markers, and the combination of both of them, as well as what the reason may be for not borrowing *nde*, or *mbe* instead of the MC.

(85) Lingala

ngái ndé (moto) nazalí koloba
 1SG FOC (MOTO) 1SG.be.PRF 15.talk
 ‘It’s me who is talking (rather than someone else).’

As a last point for further discussion, we present some data from a language that did not form part of our original selection of languages for lack of data, but which shows that *moto* may develop even further. On the road to the ultimate stage for *moto* as focus marker, a striking example is found in Ipala (J. P. Angenot, personal communication).¹⁸ Ipala is special compared to the languages we examined, not only because it displays a difference in the use of the augment or pre-prefix (cf. de Blois 1970), as already mentioned in Section 3.3, but also because the focus marker is not confined to cleft(-like) sentences.¹⁹ That is, it can be used in fragment answers to focus a preceding noun, as in (86b), which is not grammatical for the MC in the other languages discussed. This suggests that ellipsis can be an additional test for the status of the development of a focus marker (cf. Harris and Campbell 1995 and Jendraschek 2009).

(86) Ipala

- a. *ní wàlúpátá tààtá?*
 who bit father
 ‘Who bit father?’
- b. *ìlúlú múúthù àndáákō mùùthù*
 ghost MOTO not person
 ‘It’s a ghost, not a human.’
 (J.P. Angenot, personal communication)

¹⁸ The relation between the region where Ipala is spoken and the areas of Lingala and Kituba is impossible to determine with the available data, so more research would be needed on the area in between, with languages like Kimbundu, for instance.

¹⁹ Hohegger (1981: 47) mentions the use of *mntu* in Kituba as a marker of non-clefted relative clauses, but this has not been attested otherwise.

Acknowledgments: This paper is part of the research project Grammaticalization and (Inter)subjectification (GRAMIS), funded by the Belgian Science Policy (<http://webh01.ua.ac.be/gramis>). We are very grateful to Joseph Koni Muluwa (Kituba, Lingala), Léon Mundeke (Embuun), André Motingea (Ngombe, Lingala), Damase Ndembe Nsasi (Kiyoombe), André Makokila (Kimanyaanga, Kituba), André Epanga (Embudza, Lingala), Julien Mbwangi (Kituba, Lingala), Guy Ndouli (Kituba) and Joseph Djongakodi (Otetela) for discussing the data, to Tom Güldemann and Yap Foong Ha for their comments at ICHL 20 in Osaka, and to Nancy Kula and two anonymous reviewers for insightful remarks to the first draft of the paper.

References

- Aboh, Enoch, Katharina Hartmann & Malte Zimmermann. 2007. *Focus strategies in African languages: The interaction of focus and grammar in Niger-Congo and Afro-Asiatic*. Berlin & New York: Mouton de Gruyter.
- Ariel, Mira. 2008. *Pragmatics and grammar*. Cambridge: Cambridge University Press.
- Bokamba, Eyamba G. 1976. On the syntax and semantics of wh-questions in Kikongo and Kiswahili. *Studies in the Linguistic Sciences* 6(2). 65–88.
- Brinton, Laurel & Dieter Stein. 1995. Functional renewal. In Henning Andersen (ed.), *Historical linguistics 1993: Selected papers from the 11th International Conference on Historical Linguistics*, 33–47. Amsterdam & Philadelphia: John Benjamins.
- Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The evolution of grammar*. Chicago & London: The University of Chicago Press.
- Croft, William. 2000. *Explaining language change: An evolutionary approach*. Harlow & Essex: Pearson Education.
- Crowley, Terry. 1991. Parallel development and shared innovation: some developments in Central Vanuatu inflectional morphology. *Oceanic Linguistics* 30(2). 179–222.
- de Blois, Kornelis Frans. 1970. The augment in the Bantu languages. *Africana Linguistica* IV. 85–165.
- del Fabbro, Rafael & Flaviano Petterlini. 1977. *Gramática Kikongo*. Padua: Missionarios Capuchinhos em Angola.
- É. Kiss, Katalin. 1998. Identificational vs. information focus. *Language* 74(2). 245–273.
- É. Kiss, Katalin. 2006. Focussing as predication. In Valeria Molnár & Susanne Winkler (eds), *The architecture of focus*, 169–193. Berlin & New York: Mouton de Gruyter.
- É. Kiss, Katalin. 2010. Deriving the properties of structural focus. In Arndt Riester & Edgar Onea (eds.), *Focus at the syntax-semantics interface*. Working Papers of the SFB 732, vol. 3, University of Stuttgart.
- Eckardt, Regine. 2006. *Meaning change in grammaticalization*. Oxford: Oxford University Press.
- Fehderau, Harold W. 1966. *The origin and development of Kituba*. Ithaca, NY: Cornell University dissertation.
- Fiedler, Ines & Anne Schwarz. 2010. *The expression of information structure: A documentation of its diversity across Africa*. Amsterdam & Philadelphia: John Benjamins.

- Fiedler, Ines, Katharina Hartmann, Brigitte Reineke, Anne Schwarz & Malte Zimmermann. 2010. Subject focus in West African languages. In Caroline Féry & Malte Zimmermann (eds.), *Information structure: Theoretical, typological and experimental perspectives*, 234–257. Oxford: Oxford University Press.
- Fischer, Andreas. 1997. With this ring I thee wed: The verbs to wed and to marry in the history of English. In Raymond Hickey & Stanislaw Puppel (eds.), *Language history and linguistic modelling: A festschrift for Jacek Fisiak on his 60th birthday*, 467–481. Berlin & New York: Walter de Gruyter.
- Guthrie, Malcolm. 1948. *The classification of the Bantu languages*. London: Oxford University Press.
- Harris, Alice C. 2001. Focus and universal principles governing simplification of cleft structures. In Jan Terje Faarlund (ed.), *Grammatical relations in change*, 159–170. Amsterdam & Philadelphia: John Benjamins.
- Harris, Alice C. & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Hartmann, Katharina & Malte Zimmermann. 2007. In place – out of place? Focus in Hausa. In Kerstin Schwabe & Susanne Winkler (eds.), *On information structure, meaning and form*, 365–403. Amsterdam & Philadelphia: John Benjamins.
- Heine, Bernd, Ulrike Claudi & Friederike Hünemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago & London: The University of Chicago Press.
- Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd & Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske.
- Heine, Bernd, Tom Güldemann, Christa Kilian-Hatz, Donald A. Lessau, Heinz Roberg, Mathias Schladt & Thomas Stolz. 1993. Conceptual shift: A lexicon of grammaticalization processes in African languages. *Afrikanistische Arbeitspapiere* 34/35.
- Hochegger, Hermann. 1981. *Grammaire du kikongo ya leta*. Bandundu: Publications CEEBA, series III, vol. 6.
- Hopper, Paul. 1991. On some principles of grammaticalization. In Elizabeth Traugott & Bernd Heine (eds.), *Approaches to grammaticalization, vol. 1*, 17–35. Amsterdam & Philadelphia: John Benjamins.
- Hopper, Paul & Elizabeth Traugott. 2003. *Grammaticalization*, 2nd edn. Cambridge: Cambridge University Press.
- Hulstaert, Gustaaf. 1966. *Grammaire du lomongo. Troisième partie: Syntaxe*. Tervuren: MRAC, Annales Sciences humaines 58.
- Jendraschek, Gerd. 2009. Origin and development of the latmul focus construction: Subordination, desubordination, resubordination. *Folia Linguistica* 43(2). 345–390.
- Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8. 63–99.
- Kidima, Lukowa. 1990. Tone and syntax in Kiyaka. In Sharon Inkelas & Draga Zec (eds.), *The phonology-syntax connection*, 195–217. Chicago & London: The University of Chicago Press.
- Krifka, Manfred. 2007. Basic notions of information structure. In Caroline Féry, Gisbert Fanselow & Manfred Krifka (eds.), *Working papers of the SFB 632: Interdisciplinary studies on information structure (ISIS)*, 13–55. Potsdam: Universitätsverlag Potsdam.

- Kuteva, Tania. 2008. On the frills of grammaticalization. In María José López-Couso, Elena Seoane & Teresa Fanego (eds.), *Rethinking grammaticalization: New perspectives*, 189–217. Amsterdam & Philadelphia: John Benjamins.
- Lambrecht, Knud. 2000. When subjects behave like objects: An analysis of the merging of S and O in sentence-focus constructions across languages. *Studies in Language* 24(3). 611–82.
- Lambrecht, Knud. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press.
- Lichtenberk, Frantisek. 1991. On the gradualness of grammaticalization. In Elizabeth Traugott & Bernd Heine (eds.), *Approaches to grammaticalization, vol. 1*, 37–80. Amsterdam & Philadelphia: John Benjamins.
- Maho, Jouni. 2009. New Updated Guthrie List, a referential classification of the Bantu Languages. <http://goto.glocalnet.net/mahopapers/nuglonline.pdf>
- Meeuwis, Michael. 1998. *Lingala*. Munich: Lincom Europa.
- Mfoutou, Jean-Alexis. 2009. *Le français et les langues endogènes au Congo-Brazzaville*. Paris: L'Harmattan.
- Mouandza, Jean-Daniel. 2001. *Éléments de description du iyaa*. Nice: Université de Nice Sophia-Antipolis dissertation.
- Mufwene, Salikoko S. 1988. Formal Evidence of Pidginization/Creolization in Kituba. *Journal of African Languages and Linguistics* 10. 33–51.
- Mufwene, Salikoko S. 2006. Grammaticalization is part of the development of creoles. *PAPIA* 16. 5–31.
- Niyibizi, Shadrack M. 1987. *Esquisse structurale du sengele*. Brussels: Université libre de Bruxelles thesis.
- Norde, Muriel. 2002. The final stages of grammaticalization: Affixhood and beyond. In Ilse Wischer & Gabriele Diewald (eds.), *New reflections on grammaticalization*, 45–65. Amsterdam & Philadelphia: John Benjamins.
- Nsuka Nkutsi, François. 1982. *Les structures fondamentales du relatif dans les langues bantoues* (Annales Sciences Humaines 108). Tervuren: Musée Royal de l'Afrique Centrale.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75–116.
- Rooth, Mats. 1996. Focus. In Shalom Leppin (ed.), *The handbook of contemporary semantic theory*, 271–297. Oxford: Blackwell.
- Sabel, Joachim & Jochen Zeller. 2006. WH question formation in Nguni. In John Mugane, John P. Hutchison & Dee A. Worman (eds.), *Selected proceedings of the 35th Annual Conference on African Linguistics*, 271–283. Somerville, MA: Cascadilla Proceedings Project.
- Samarin, William. 1991. The Origins of Kituba and Lingala. *Journal of African Languages and Linguistics* 12. 47–77.
- Selkirk, Elizabeth O. 1984. Phonology and syntax. *The relation between sounds and structure*. Cambridge, MA: MIT Press.
- Selkirk, Elizabeth O. 1995. Sentence prosody: Intonation, stress and phrasing. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 550–569. Oxford: Blackwell.
- Tete Wer Sey, K. 1975. *Éléments de grammaire wongo (wongk) C85, phonologie et morphologie*. Lubumbashi: Université Nationale du Zaïre thesis.
- Thomason, Sarah G. 2001. Contact-Induced Language Change and Pidgin/Creole Genesis. In Norval Smith & Tonjes Veenstra (eds.), *Creolization and contact*, 249–262. Amsterdam & Philadelphia: John Benjamins.

- Toronzoni, Ngama-Nzombio. 2004. *Esquisse du Bomboma (langue bantu de zone C)*. Kinshasa: Centre de Recherche de pédagogie appliquée.
- Tosco, Mauro. 2002. "A whole lotta focusin' goin' on": Information packaging in Somali texts. *Studies in African Linguistics* 31(1/2). 27–53.
- Traugott, Elizabeth Closs. 1988. Pragmatic strengthening and grammaticalization. *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*. 406–416.
- Traugott, Elizabeth Closs. 1995. The role of the development of discourse markers in a theory of grammaticalization, handout paper ICHL XII, version 1997. <http://www.stanford.edu/~traugott/papers/discourse.pdf>.
- Traugott, Elizabeth & Graeme Trousdale. 2010. Gradience, gradualness and grammaticalization: How do they intersect? In Elizabeth Traugott & Graeme Trousdale (eds.), *Gradience, gradualness and grammaticalization*, 19–44. Amsterdam & Philadelphia: John Benjamins.
- van der Wal, Jenneke. 2009. Word order and information structure in Makhuwa-Enahara. Utrecht: LOT.
- van der Wal, Jenneke. 2011. Focus excluding alternatives: conjoint/disjoint marking in Makhuwa. *Lingua* 121(11). 1734–1750.
- van der Wal, Jenneke. forthcoming. Diagnosing focus. *Graz Working Papers in Linguistics*.
- Vincent, Nigel. 1995. Exaptation and grammaticalization. In Henning Andersen (ed.), *Historical linguistics 1993: Selected papers from the 11th International Conference on Historical Linguistics*, 433–448. Amsterdam & Philadelphia: John Benjamins.
- Willis, David. 2010. Degrammaticalization and obsolescent morphology. In Ekaterini Stathi, Elke Gehweikler & Ekkehard König (eds.), *Grammaticalization: Current views and issues*, 151–178. Amsterdam & Philadelphia: John Benjamins.
- Zerbian, Sabine. 2006. Expression of information structure in Northern Sotho. Berlin: Humboldt University dissertation.
- Zerbian, Sabine. 2007. The subject/object asymmetry in Northern Sotho. In Kerstin Schwabe & Susanne Winkler (eds.), *On information structure, meaning and form: Generalizations across languages*, 323–364. Amsterdam & Philadelphia: John Benjamins.

Appendix. Lingala and Kikongo areas

