

## prefinal version

Van de Velde, Mark (2019). Nominal morphology and syntax. In: Van de Velde, Mark; Koen Bostoen; Derek Nurse & Gérard Philippson (eds). *The Bantu Languages*, 2nd edition [Routledge Language Family Series], 237-269. London: Routledge.

## CHAPTER EIGHT

# Nominal Morphology and Syntax

*Mark Van de Velde*

## 1 INTRODUCTION

One of the quintessential typological properties of the Bantu languages is their pervasive system of noun classes and noun class agreement. This is undoubtedly the aspect of their grammatical structure that is most discussed in the literature, if only because every grammar sketch of a Bantu language contains a section on noun classes. The most complete discussion can be found in Maho (1999). In contrast, the structure of the noun phrase has received little attention, which cannot be attributed to a lack of interesting features. This chapter briefly introduces aspects of the noun and noun phrase that are well studied, such as the noun class system and the different types of adnominal modifiers, and provides them with a reference. It also aims at filling some gaps in the literature. Section 2.2.2.2 provides a first systematic overview of the types of semantic agreement that can be found in the Bantu languages. Section 3 is entirely dedicated to the augment, a pervasive element in the grammar of the Bantu languages, of which the last comparative study (De Blois 1970) was in need of an update. The typologically unusual word order patterns that can be found in the Bantu languages receive a first comparative analysis and diachronic explanation in Section 5.

## 2 NOUNS AND NOUN CLASSES

### 2.1 The form and structure of nouns

Bantu nouns minimally consist of a stem, which is usually preceded by a class prefix, itself sometimes preceded by one or more prefixes or proclitics, which can be inflectional or derivational. Noun stems can contain a derivational suffix (cf. Schadeberg & Bostoen, this volume). Reconstructed Proto-Bantu noun stems tend to have a CVCV-structure. In many contemporary languages, this is still the canonical stem form. However, in some north-western languages, such as Nzadi B865 (Crane *et al.* 2011) and Nsong B85d (cf. Koni Muluwa & Bostoen this volume), the last vowel or syllable of noun stems tends to have

dropped, giving rise to a monosyllabic pattern as the most frequent stem type. The prefix + stem template is so strong in the Bantu languages, that monomorphemic nouns may behave prosodically as if they consisted of a prefix and a stem, and that class prefixes that were historically incorporated in stems can be reactivated in different ways.

As for tone, augments and derivational prefixes added to full nouns, i.e. class prefix plus stem, are reconstructed high, except in class 1 and 9, whereas class prefixes are reconstructed low (Meeussen 1967). Proto-Bantu disyllabic stems can have any of the four logically possible tone patterns, viz. \*HH, \*HL, \*LL and \*LH (Bastin *et al.* 2002). Here too, many changes have taken place in individual languages, such as the appearance of H class prefixes in a subset of Punu B43 nouns (Blanchon 1997) or the emergence of two extra tone patterns on disyllabic noun stems in Nande JD42 (Kenstowicz 2008). In a number of western Bantu languages, the tone pattern of nouns depends on their syntactic position, a phenomenon known as ‘tone case’ in Bantu studies (see for instance Schadeberg 1986, Kavari *et al.* 2012).

## 2.2 The noun class system

Noun classes can be defined as sets of nouns that trigger the same agreement pattern. Noun class assignment is typically coded by means of a nominal class prefix in the Bantu languages. Bantuists use numbers, rather than labels such as ‘feminine’ or ‘neuter’, to refer to individual noun classes. Numbers are assigned to classes in individual languages on the basis of cognacy. Odd numbers are used for classes that contain singular nouns and even numbers for plural classes, with some exceptions, most notably class 12 (SG) and 13 (PL). Singular – plural class pairings are usually called ‘genders’. The classic Swahili G42 examples in (1) show the noun class prefixes and agreement prefixes of the noun *-kapu* ‘basket’ in the singular (class 7) and plural (class 8).

- (1) Swahili G42 (Welmers 1973: 171)
- a. *Ki-kapu ki-kubwa ki-moja ki-li-anguk-a.*  
 7-basket NP<sub>7</sub>-big NP<sub>7</sub>-one SP<sub>7</sub>-PST-fall-FV  
 ‘One large basket fell.’
- b. *Vi-kapu vi-kubwa vi-tatu vi-li-anguk-a.*  
 8-basket NP<sub>8</sub>-big NP<sub>8</sub>-three SP<sub>8</sub>-PST-fall-FV  
 ‘Three large baskets fell.’

Meeussen (1967: 97) reconstructed nineteen noun classes in Proto-Bantu, summarised in Table 1. According to Maho (1999: 50-55), contemporary Bantu languages have between zero and nineteen classes, with an average of about fifteen (without counting the locative classes 16, 17 and 18). The most common genders are 1/2, 3/4, 5/6, 7/8, 9/10, 11/6, 11/10, 12/13 and 14/6 (Maho 1999: 54). Classes from number 12 (included) upward generally contain less nouns than the other classes. The distribution of nouns over the noun classes is little studied, but appears to be quite variable across Bantu. Class 1, for instance, tends to be very small in (north)western languages and larger in the East and South. Class 7, in contrast, is usually much larger in the West than in the East.<sup>1</sup>

	Nom	Num	Pron	Subj	Obj
1	<i>mù-</i>	<i>(ù?)</i>	<i>jù-</i>	<i>ú-, á-</i>	<i>mù-</i>
2	<i>bà-</i>	<i>bá-</i>	<i>bá-</i>	<i>bá-</i>	<i>bá-</i>
3	<i>mù-</i>	<i>(ú?)</i>	<i>gú-</i>	<i>gú-</i>	<i>gú-</i>

4	<i>mì-</i>	<i>(ì-?)</i>	<i>gí-</i>	<i>gí-</i>	<i>gí-</i>
5	<i>ì-</i>	<i>dí-</i>	<i>dí-</i>	<i>dí-</i>	<i>dí-</i>
6	<i>mà-</i>	<i>(á-?)</i>	<i>gá-</i>	<i>gá-</i>	<i>gá-</i>
7	<i>kì-</i>	<i>kí-</i>	<i>kí-</i>	<i>kí-</i>	<i>kí-</i>
8	<i>bì-</i>	<i>bí-</i>	<i>bí-</i>	<i>bí-</i>	<i>bí-</i>
9	<i>n-</i>	<i>(ì-)</i>	<i>jì-</i>	<i>jí-</i>	<i>jí-</i>
10	<i>n-</i>	<i>í-</i>	<i>jí-</i>	<i>jí-</i>	<i>jí-</i>
11	<i>dù-</i>	<i>dú-</i>	<i>dú-</i>	<i>dú-</i>	<i>dú-</i>
12	<i>kà-</i>	<i>ká-</i>	<i>ká-</i>	<i>ká-</i>	<i>ká-</i>
13	<i>tù-</i>	<i>tú-</i>	<i>tú-</i>	<i>tú-</i>	<i>tú-</i>
14	<i>bù-</i>	<i>bú-</i>	<i>bú-</i>	<i>bú-</i>	<i>bú-</i>
15	<i>kù-</i>	<i>kú-</i>	<i>kú-</i>	<i>kú-</i>	<i>kú-</i>
16	<i>pà-</i>	<i>pá-</i>	<i>pá-</i>	<i>pá-</i>	<i>pá-</i>
17	<i>kù-</i>	<i>kú-</i>	<i>kú-</i>	<i>kú-</i>	<i>kú-</i>
18	<i>mù-</i>	<i>mú-</i>	<i>mú-</i>	<i>mú-</i>	<i>mú-</i>
19	<i>pì-</i>	<i>pí-</i>	<i>pí-</i>	<i>pí-</i>	<i>pí-</i>

**Table 1:** The Proto-Bantu noun class system as reconstructed by Meeussen (1967: 97)

Languages that have fully or nearly lost their noun class system are very rare among the Bantu languages, an observation that has sometimes been used as an - obviously imperfect - criterion for distinguishing Narrow Bantu from related language groups, where typologically diverse reduced class systems are common (see Good 2012 for an overview of most of Niger-Congo). An example of a language classified as Narrow Bantu that has lost its noun class system is Polri A92 (Wega Simeu 2012). A small set of about five nouns with a human denotation and a vowel initial stem has a different prefix in the singular (*mV-*) as in (2a) and the plural (*bò-*) as in (2b). Plural formation for all other nouns, such as *sámá* ‘sheep’ in (2c), involves the use of the plural marker *bè* in NP-initial position, as in in (2d). There is no agreement in number or class anywhere. Plural NPs with a prenominal modifier have to start with the plural morpheme, also when their head noun has plural marking, as shown in (2b).

(2) Polri A92 (Wega Simeu 2012)

- a. *jíké mw-ǎn jà wàngò-ím*  
other SG-child DEM come-HOD.PST  
‘That other child has come.’
- b. *bè jíké bw-ǎn jà wàngò-ím*  
PL other PL-child DEM come-HOD.PST  
‘Those other children have come.’
- c. *jíké sámá jà wàngò-ím*  
other sheep DEM come-HOD.PST  
‘That other sheep has come.’
- d. *bè jíké sámá jà wàngò-ím*  
PL other sheep DEM come-HOD.PST  
‘Those other sheep have come.’

Most Bantu languages that lack noun classes, such as Polri A92 or Bila D32, are in contact with non-Benue-Congo languages, and contact induced change has been suggested as an explanation for the loss of classes in these cases (see for instance Dimmendaal 2011: 196).

However, Nzadi B865 shows that the very strong reduction of noun classes can also be due to purely language-internal changes (Crane *et al.* 2011).

### 2.2.1 Counting noun classes

Establishing the number of noun classes in a language is not as straightforward as it may seem. Different authors have used different criteria, depending on their goals. One option is to distinguish noun classes (sets of nouns that trigger the same agreement pattern) from morphological classes (sets of nouns with identical number marking). The advantages of this approach are that it provides the most accurate synchronic analysis of the nominal morphosyntax, and that it allows for straightforward cross-linguistic comparison with other systems of grammatical gender. The disadvantage of this approach is that it is not very compatible with the comparative concepts ‘class 1’, ‘class 2’, etc. developed in the Bantuist linguistic tradition, so that adopting it complicates Bantu internal comparison. The reason for this is that there are many instances of noun classes whose agreement patterns have merged in the history of individual languages, whereas their nominal class markers have remained unchanged. Therefore, most Bantuists prefer to follow the tradition of splitting up sets of nouns that trigger the same agreement pattern into two or more classes if they have different nominal prefixes AND if they participate in different singular-plural pairings, or if one set contains singular nouns and the other plural nouns. For instance, Makwe G402 has two sets of nouns that trigger the same agreement pattern in the singular: one with a prefix *mu-* and plurals of class 4; and one with a prefix *u-* ~ *lu-* and plurals in class 4, 6, 10, 10a or 10b. Following Bantuist practice, Devos (2008: 43) treats these as two separate classes, 3 and 11 respectively, which reflects the historical origin of the difference in nominal prefix.

Bantuists also have the somewhat inconsistent habit of distinguishing between a class 15 and a class 17, even though these two classes tend to have exactly the same nominal prefix and agreement pattern. Class 17 is one of the locative classes, together with 16 and 18. Its class marker can be added as a pre-prefix to a full noun to derive a locative noun, which can trigger agreement according to the locative class or its original noun class, depending on the language and/or the construction. Class 15, on the other hand, is the class of infinitives in most Bantu languages, except in the north of the domain, where infinitives are often of class 5, sometimes also other classes (cf. Forges 1983, Hadermann 1999). Class 15 also contains a small set of canonical nouns, mostly body part terms. Six have been reconstructed in Proto-Bantu with a plural in class 6: *\*-bókò* ‘arm’, *\*-gùdù* ‘leg’, *\*-túi* ‘ear’, *\*-dúi* ‘knee’ and *\*-jápà* ‘armpit’ (Doneux 1967). The arguments for splitting agreement class 15-17 into two distinct noun classes, when explicitly mentioned, are rather diverse and usually not compatible with any working definition of noun classes.

When sets of singular nouns trigger the same agreement pattern and have their plural in the same class, but differ in the shape of their nominal prefix in ways that are not phonologically predictable, they tend to be divided into subclasses.<sup>ii</sup> Bantu subclasses are typically labelled by means of a letter after the class number. For example, Devos (2008) distinguishes in Makwe G402 between class 10 (prefix *ji-*), class 10a (prefix  $\emptyset$ -) and class 10b (prefix *jiN-*). Such subclasses and their labels are typically language-specific, or even description-specific, with one notable exception: class 1a, which can be found throughout Bantu. The so-called class 1a was first systematically described by Doke (1927) as a set of nouns that lack a prefix in the singular and usually trigger class 1 agreement. It typically contains proper names, some kinship terms, personified animals and borrowings. Class 1a is radically different from the other subclasses. Its lack of a nominal prefix in most languages is not due to prefix loss, but

goes back to Proto-Bantu at least. It can be explained by pointing out that proper names and certain suppletive kin terms tend to lack a determiner in the languages of the world, or an augment in the Bantu languages (see Section 3), because they are inherently determined. Names for personified animals function as proper names and borrowings are easily attracted to any class without an overt class marker. The plurals of class 1a nouns are often marked by an element *ba(a)* or *(b)ɔ* in which case they trigger the same agreement pattern as nouns of class 2 and are treated as belonging to subclass 2a, when the vowel of the marker is *a*, or 2b, when the plural marker has a back vowel. In Eton A71, the class 2b marker *bɔ* has the phonological characteristics of a separate word and I have analysed it as a number word (Van de Velde 2006a), an analysis that may be extended to other Bantu languages. Class 2a/2b markers often express associativity. Associative constructions are used to refer to collectives that consist of an identified referent and a number of associate referents. The relation between the identified referent and its associates can be metonymic, i.e. based on contiguity, or metaphoric, i.e. based on similarities. The former type is typically reserved for human referents in the languages of the world (3a), but in some Bantu languages it can also be used with place names (4a). Examples of metaphoric associativity, also called ‘similitive plurals’ (Daniel & Moravcsik 2005), can be found in different parts of the Bantu speaking area, as shown in (3b) and (4b).

- (3) Mongo C61 (Hulstaert 1965: 145)
- a. *baa Byeka*  
‘Byeka and his family/pupils/followers...’
  - b. *baa mésá*  
‘tables and similar things; tables, for instance; tables and the like’
- (4) Xhosa S41 (Hendrikse 1990: 391), *oo* is the class 2b marker
- a. *oomaRhini* < *oo* + *amaRhini*  
‘Grahamstown and environment’
  - b. *ookulamba* < *oo* + *ukulamba*  
‘hunger and similar feelings; hunger and the like’

### 2.2.2 Semantics

There is widespread agreement among Bantuists that the gender assignment of a noun cannot generally be predicted on the basis of its meaning, but that the noun class systems of Bantu languages are not devoid of semantic regularities either. Disagreement exists about whether or not the gender assignment of all or most nouns can be shown to be semantically motivated, if not in contemporary languages, then at least in a proto-stage.

One area in which gender assignment could in principle be clearly shown to be semantically motivated is the integration of borrowings into the noun class system (cf. Mous, this volume). Unfortunately, there is no thorough comparative study on this subject. In my experience, gender assignment on semantic grounds is relatively exceptional and mostly restricted to language names and some nouns for human beings. Most often, borrowings are assigned to classes on formal grounds. Either they are inserted into classes that have no, or no clearly recognisable prefix, such as class 9, 5 or 1a, or the initial segments of their stem are reinterpreted as an existing class prefix. Note in this respect that for comparative-historical reasons class 9 is often analysed as having a nasal prefix *N-*, but that in many languages the plural (class 10) has the same initial nasal, which therefore does not commute with anything

else, and is strictly speaking not morphologically separable from the stem, unless, perhaps, in some deverbal nouns, where it can be analysed as a derivational affix.

The remainder of this section will be devoted to a number of uncontroversial semantic characteristics of Bantu class systems, starting with gender assignment and then discussing meaning in grammar. The human/non-human distinction will be a recurrent theme. Due to space limitations, this section cannot do justice to the extensive literature on the subject. A more elaborate introduction with many references can be found in Maho (1999: 63-99).

#### 2.2.2.1 Gender assignment

The first semantic regularity that naturally comes to mind is the semantic homogeneity of gender 1/2 in many languages, which tends to be restricted to nouns with human reference. However, it is quite common to find terms for humans in other classes, especially if they are used to refer to persons with unusual characteristics. Second, words for liquids are typically found in class 6. Other semantic sub-regularities in gender assignment that have been pointed out in the literature are not Bantu-wide, or they tend to show many more exceptions. Language names, for instance, tend to be of class 7 or of class 11 and nouns for abstract concepts are in class 14 in many languages.

Many patterns of nominal derivation have a target gender, in which they create semantically coherent subgroups. Agentive nouns assigned to gender 1/2 are a typical example (cf. Schadeberg & Bostoen, this volume). Special cases are evaluative and locative denominal derivation. When used derivationally, locative class markers are normally additive, i.e. added to the inherent class marker, whereas diminutive and augmentative prefixes can be either substitutive, i.e. replacing the inherent class marker, or additive, with a general preference for the former. Typical diminutive classes are 12 and 19, sometimes 5. Class 7 is the most recurrent augmentative class. Other combinations of classes and evaluative meanings can be found too. For instance, class 5 can have an augmentative meaning in some languages and class 7 a diminutive one.

The grammaticalisation of evaluative morphemes can lead to the emergence of new class markers. In Eton A71, a diminutive can be formed of any noun by means of the proclitic  $m\dot{\delta}^H =$  (plural  $b\dot{\delta}^H =$ ), historically derived from the noun  $m-\dot{\delta}\eta\acute{\delta}$  (cl.1) /  $b-\dot{\delta}\eta\acute{\delta}$  (cl.2) 'child(ren)'. These procliticised diminutives trigger agreement pattern 1 in the singular and 2 in the plural, as predicted by their etymology. Leitch (2003: 400) analyses Babole C101 as having a diminutive gender 19/20, with a prefix  $mw\acute{a}-$  in the singular and  $b\acute{a}n\acute{a}-$  in the plural. These too are derived from the gender 1/2 noun 'child', but they can be analysed as the markers of a separate noun class, because they trigger a unique agreement pattern. These agreement patterns correspond to those of class 19 and 13 respectively in other Bantu languages, typical diminutive classes. So at one point in the history of Babole, the noun for 'child' must have been used as the head of a genitive phrase to express a diminutive meaning, in which case it triggered semantic agreement.

#### 2.2.2.2 Meaning and grammar

We speak of semantic agreement, as opposed to syntactic agreement, when the choice of an agreement pattern depends on aspects of the meaning of the controller rather than on its morphological class defined by the nominal prefix (Corbett 1991). Here, the link between noun classes and semantics is more straightforward than in gender assignment. The best

known example is animate agreement, as found especially in zone K and among the coastal languages of zone G and E (Wald 1975, Maho 1999: 124), where animate nouns trigger agreement pattern 1 in the singular and 2 in the plural, whatever the morphological class to which they belong, as shown in (5b), where the Swahili class 7 noun *kiboko* ‘hippo’ triggers class 1 agreement of both the demonstrative and the verbal object prefix, in contrast to *kisu* ‘knife’ which triggers syntactic class 7 agreement.

(3) Swahili G42 (Wald 1975: 241-242)

- |    |                         |                 |                       |
|----|-------------------------|-----------------|-----------------------|
| a. | <i>ki-le</i>            | <i>ki-su,</i>   | <i>ni-li-ki-on-a.</i> |
|    | PP7-DEM                 | 7-knife         | SP1SG-PST-OP7-see-FV  |
|    | ‘That knife, I saw it.’ |                 |                       |
| b. | <i>Yu-le</i>            | <i>ki-boko,</i> | <i>ni-li-mw-ona.</i>  |
|    | PP1-DEM                 | 7-hippo         | SP1SG-PST-OP1-see-FV  |
|    | ‘That hippo, I saw it.’ |                 |                       |

There is some variation between languages concerning the obligatoriness of its application and the agreement targets that are involved. For the latter, Corbett’s (1979) agreement hierarchy (attributive < predicate < relative pronoun < personal pronoun) makes the right predictions, viz. the higher (i.e. the more to the right) an agreement target is on the hierarchy, the more likely it is to have semantic agreement. In fact, the Bantu languages might require a refinement of the hierarchy, since adnominal modifiers can be split into possessive pronouns versus the others. Possessive pronouns are at the bottom of the hierarchy, i.e. most likely to agree according to the morphological class of the head noun (more on this in Section 5).

The notion of animate agreement can only make sense when it is opposed to syntactic agreement. It should not be confounded with the situation found in the few Bantu languages that have a radically restructured class system in which animacy determines gender assignment. An interesting example is Kako A93, which has the typologically unusual characteristic of distinguishing two genders in the plural, but none in the singular (Ernst 1992: 34). Kako does not have nominal prefixes on singular nouns. Plural nouns take the prefix *bè-* (6a-b) if they are animate and *mè-* if they are inanimate (6c-d). These prefixes also mark agreement in gender within the noun phrase (6e-f).

(4) Kako A93 (Ernst 1992)

- a. *mbam* ‘the man’ / *bè-mbam* ‘the men’
- b. *mbiyè* ‘the dog’ / *bè-mbiyè* ‘the dogs’
- c. *gwâlò* ‘the hoe’ / *mè-gwâlò* ‘the hoes’
- d. *tù* ‘the house’ / *mè-tù* ‘the houses’
- e. *bè-ngo ba-ka* ‘these pigs’
- f. *mè-kandò ma-ka* ‘these clothes’

The main variety of Lingala C36d, also called ‘Kinshasa Lingala’ (Meeuwis 2010: 37), has retained a typical Bantu system of morphological classes: eleven noun prefixes define fifteen classes plus two subclasses, if one counts in a traditional Bantuist way, taking into account singular-plural pairings of nouns. The nominal prefixes have inflectional uses (number marking) and derivational uses. However, the morphological class of a noun is completely irrelevant for the agreement it triggers. Agreement in gender is restricted to subject markers on verbs and defines two genders: Animate and Inanimate. Animate controllers trigger a subject prefix *a-* in the singular and *ba-* in the plural; inanimate controllers trigger the subject marker *e-* in the singular and the plural. Therefore, despite its full set of nominal prefixes, the

grammatical gender of Lingala C36d is typologically more similar to that of Kako A93 or English than to that of the majority of Bantu languages.

	SG	PL
1/2 1a	<i>mo-</i> $\emptyset$ -	<i>ba-</i>
3/4	<i>mo-</i>	<i>mi-</i>
5/6	<i>li-</i>	<i>ma-</i>
7/8 7a	<i>e-</i> <i>ki-</i>	<i>bi-</i>
9/10	$\emptyset$ -	<i>ba-</i>
11	<i>lo-</i>	
14	<i>bo-</i>	
15	<i>ko-</i>	

**Table 2:** Morphological classes in Lingala C36d

Another kind of semantic agreement could be called ‘superclassing’, in that it involves a hierarchical organisation of the agreement patterns of a language along semantic lines. This is described for a number of multi-class stems in Ciluba L31a, such as the non-selective interrogative stem *-nyi*, the indefinite stem *-ntu* and the locative adverbials (van den Eynde & Mufuta 1994). Thus, wherever one can use the indefinites *mu-ntu* ‘somebody’ (class 1), *ci-ntu* ‘something’ (class 7) or *ka-ntu* ‘a small person or thing’ (class 12), the use of the more general term *ci-ntu* ‘some entity’ (class 7) is felicitous as well. Likewise, wherever one can use *ku-ntu* ‘some direction’ (class 17) or *mu-ntu* ‘some place inside’ (class 18), one can also always use *pa-ntu* ‘somewhere’ (class 16), which can have the more specific meaning of ‘some place outside’ too. This shows class 7 and 16 to be at the top of a semantic hierarchy of agreement patterns for entities and locations respectively. The locative hierarchy is illustrated by the use of adverbials in (7) (van den Eynde & Mufuta 1994: 102). In examples (7a) and (7b) the same message can be expressed in two ways. The initial demonstrative can either agree in noun class with the locative noun, or it can take the default locative form of class 16. These options are not available in (7c), because the default locative class is identical to the one for the specific meaning of superposition. Examples (7d-g) are ungrammatical because the demonstrative fails to agree with the noun, whether by syntactic agreement or by superclassing.

(5) Ciluba L31a (van den Eynde & Mufuta 1994)

- a. *apa tu-di ku-n-zubu* or *eku tu-di ku-n-zubu*  
DEM<sub>16</sub> SP<sub>1PL</sub>-be 17-9-house DEM<sub>17</sub> SP<sub>1PL</sub>-be 17-9-house  
‘Here, we are towards the house’
- b. *apa tu-di mu-n-zubu* or *emu tu-di mu-n-zubu*  
DEM<sub>16</sub> SP<sub>1PL</sub>-be 18-9-house DEM<sub>18</sub> SP<sub>1PL</sub>-be 18-9-house  
‘Here, we are in the house’
- c. *apa tu-di pa-n-zubu*  
DEM<sub>16</sub> SP<sub>1PL</sub>-be 16-9-house  
‘Here, we are on the house’
- d. \**emu tu-di ku-n-zubu*
- e. \**emu tu-di pa-n-zubu*
- f. \**eku tu-di mu-n-zubu*



g. \**eku tu-di pa-n-zubu*

A similar semantically based hierarchical organisation of the class system is at work in contexts of enforced agreement, e.g. where conjoined NPs in subject position have to agree with the verb, a situation also known as concord resolution (cf. Marten & Downing, this volume, Section 3.4).

A special case of enforced agreement can be found with proper name and suppletive kin term controllers (Van de Velde 2006b, 2009). In some languages they trigger a default agreement pattern, e.g. class 9 for place names in Makwe G402 (Devos 2008: 61). This default pattern may be mixed, as in Orungu B11b, where names trigger class 9 agreement within the nominal constituent and class 1 agreement on verbs and pronouns (Van de Velde & Ambouroué 2011). Similarly, in Kagulu G12, suppletive kin terms trigger agreement pattern 5 (SG) or 10 (PL) on possessive pronouns and 1/2 on other agreement targets (Petzell 2008: 56). In other languages, such as Gikuyu E51 and Rundi JD62, as shown in (8), proper names and suppletive kin terms trigger the same agreement pattern as the basic level term that expresses their categorical presuppositional meaning (Van de Velde 2009).

(6) Rundi JD62 (Van de Velde 2009)

a. *u-ru-kara* ‘black’ (cl. 11); *u-muu-ntu* ‘person’ (cl.1); *i-m-bwá* ‘dog’ (cl.9)

b. *Rukara a-rikó a-ra-fuungur-a*  
Rukara SP<sub>1</sub>-PROG SP<sub>1</sub>-DJ-eat-IPFV  
‘Rukara (a person) is eating.’

c. *Rukara i-rikó i-ra-ry-á*  
Rukara SP<sub>9</sub>-PROG SP<sub>9</sub>-DJ-eat-IPFV  
‘Rukara (a dog) is eating.’

d. *u-muu-ntu* ‘person’ (cl. 1); *i-nká* ‘cow’ (cl. 9)

e. *nyina a-ra-ryam-ye*  
mother SP<sub>1</sub>-DJ-sleep-PFV  
‘His/her mother is sleeping.’ (person)

f. *nyina i-ra-ryam-ye*  
mother SP<sub>1</sub>-DJ-sleep-PFV  
‘His/her mother is sleeping.’ (cow)

The so-called class 1a, the mixed agreement patterns of Orungu B11b and Kagulu G12 and the semantic proper name agreement of Rundi JD62 and Gikuyu E51 are basically the same phenomenon, which can be dealt with coherently by analysing proper names and grammatically similar common nouns as being classless and the agreement they trigger as a kind of enforced agreement (Van de Velde 2006b).

So far, we have seen three types of semantic agreement, viz. animate agreement, superclassing and agreement determined by the categorical sense of proper name controllers. A fourth type could be called ‘evaluative agreement’. It can be observed in at least two contexts. The first is found in those languages with animate agreement, such as Swahili G42 and Ndengeleko P11 (Ström 2013: 190), where nouns with animate reference obligatorily trigger agreement patterns 1 or 2 on all or most targets, except when they are derived diminutive or augmentative nouns. In that case, as illustrated in (9), they trigger the agreement pattern of the diminutive or augmentative class.

(7) Ndengeleko P11 (Ström 2013: 163, 195)

- a. *m-bésa a-úu*  
10-hare NP<sub>2</sub>-white  
'White hares.'
- b. *ka-pésa ka-úu*  
12-hare NP<sub>2</sub>-white  
'Little white hare.'

When the lexical class of a noun with animate reference is the same as one of the evaluative classes, as in Swahili G42 (10), agreement alone signals whether the noun is derived (Gregersen 1967: 19). Adherents of modular approaches to grammar may find this an interesting example of denominal derivation signalled by the syntax rather than the morphology.

(8) Swahili G42

- a. *Yu-le ki-pofu, ni-li-mw-on-a*  
PP<sub>1</sub>-DEM 7-blind SP<sub>1SG</sub>-PST-OP<sub>1</sub>-see-FV  
'That blind man, I saw him.'
- b. *Ki-le ki-pofu, ni-li-ki-on-a*  
PP<sub>7</sub>-DEM 7-blind 1SG-PST-OP<sub>7</sub>-see-FV  
'That tiny blind man, I saw him.'

Agreement in examples like (8b) and (8e) is usually analysed as syntactic agreement, e.g. in Wald (1975: 273). However, it is impossible to account for it without making reference to the semantic properties of the head noun. The fact that evaluative agreement outranks animate agreement can be ascribed to the conspiracy of formal and semantic conditions in evaluative agreement.

The second context in which evaluative agreement can be observed provides independent evidence for its analysis as a type of semantic agreement. It is found in languages such as Rundi JD62, where the agreement pattern of proper name and suppletive kin term controllers is determined by their presuppositional meaning. Since emotive connotations are typically part of the meaning of names (Van Langendonck 2007: 83), it is not surprising that they can trigger evaluative agreement of diminutive class 12 or augmentative class 7, as shown in (11a-c). This option is not available for common noun controllers, which always trigger syntactic agreement, as shown in (11d-e).

(9) Rundi JD62 (Van de Velde 2009)

- a. *Taama a-ra-aje*  
Taama SP<sub>1</sub>-DJ-come.PFV  
'Taama arrives'
- b. *Taama ki-ra-aje*  
Taama SP<sub>7</sub>-DJ-come.PFV  
'(big/horrible) Taama arrives.'
- c. *Taama ka-ra-aje*  
Taama SP<sub>12</sub>-DJ-come.PFV  
'(little/dear) Taama arrives.'
- d. *u-mu-ganwa a-ra-aje*  
AUG-1-prince SP<sub>1</sub>-DJ-come.PFV  
'The prince arrives.'
- e. *\*u-mu-ganwa ka-ra-aje*

AUG-1-prince SP<sub>12</sub>-DJ-come.PFV  
'The (little/dear) prince arrives.'

The fifth and last type of semantic agreement is locative agreement. It can be argued to exist in some of the languages that derive locative nouns from other nouns by means of a suffix. These languages are spoken in zones E, G, F, P and S, i.e. the Eastern part of the Bantu area from North to South. The most common locative suffix is a form similar to *-ini* (Grégoire 1975). There is a small group of languages of zone G in which locative nouns are derived from non-locative nouns uniquely by means of a suffix, i.e. without also adding a locative class prefix, and in which locative nouns can trigger agreement of class 16, 17 or 18 according to the type of location the speaker wishes to express, as the Bondei G24 examples in (12) show. Even if one chooses to analyse the locative suffix as a formally unusual class marker, the choice between three different agreement patterns can only be analysed in terms of semantic agreement.

- (10) Bondei G24 (Grégoire 1975: 192)
- |    |                         |                           |                      |
|----|-------------------------|---------------------------|----------------------|
| a. | <i>nyumba-ni mw-ako</i> | (18-POSS <sub>2SG</sub> ) | 'in your house'      |
| b. | <i>nyumba-ni ha-kwe</i> | (16-POSS <sub>3SG</sub> ) | 'close to his house' |
| c. | <i>nyumba-ni kw-etu</i> | (17-POSS <sub>1PL</sub> ) | 'at our house'       |

To conclude, although lexical gender assignment has received most attention in studies on the semantic basis of Bantu noun class systems, evidence for semantic regularities is stronger in the grammar, i.e. derivation and agreement.

### 3 THE AUGMENT

The term 'augment' (also 'initial vowel' or 'pre-prefix') is used by Bantuists to refer to an element that precedes the class prefix of nouns and some adnominal or nominalised modifiers and that changes neither the class assignment of the noun, nor its lexical meaning, nor the syntactic positions it can occupy. In Meeussen's (1967) tentative Proto-Bantu reconstruction, the augment is a separate word that is formally identical to a pronominal prefix. Since the augment is pervasive in Bantu and since it hasn't been the object of a comparative study since De Blois (1970), I will go into considerable detail discussing the shape (Section 3.1), use (Section 3.2) and origin (Section 3.3) of the augment.

#### 3.1 Shape

Contemporary languages can be typologised according to the shape of their augment (cf. De Blois 1970: for a comparative overview). The reconstructed CV-shape, identical to the reconstructed shape of the pronominal prefix, can be found in a very restricted set of Eastern languages of zones J and G. Interestingly, the augment in these languages often only has a CV-shape in classes 1, 3, 4, 6 and 10, i.e. the nasal classes, in which the segmental form of the pronominal prefix differs from that of the nominal prefix. In other classes, the form of the augment is reduced to a vowel identical to that of the class prefix. A variation on this pattern can be found in Swati S43, for example, where the nasal classes have a V-augment and the other classes have no segmental augment (Ziervogel & Mabuza 1976). Given its distribution, the reduction of the augment to a single vowel may be explainable as partial haplology (see Patin et al. this volume for a synchronic case of allomorphy of the augment of class 7 that can be interpreted along the same lines). In most languages that have an augment, its shape is V-

in all classes. Here, languages differ according to whether they have a “full” paradigm of three augment vowels, i.e. the three vowels that can be found in pronominal prefixes, or whether the paradigm has been partly or fully reduced. An example of the former is Rundi JD62 in (13), where we find *i-* in class 4, 5, 7, 8, 9 and 10, *u-* in class 1, 3, 11, 13, 14, 15 and 18, and *a-* in class 2, 6, 12 and 16 (Meeussen 1959: 62). Note that *u* and *i* are the reflexes of PB *\*u* and *\*i* in Rundi and that the augment vowels correspond to those of the pronominal prefixes reconstructed for their respective classes (see Table 1).

- (11) Rundi JD62 (Meeussen 1959: 71)
- |                  |   |                  |            |           |
|------------------|---|------------------|------------|-----------|
| <i>ù-mùù-ntù</i> | / | <i>à-bàà-ntù</i> | ‘person/s’ | class 1/2 |
| <i>ì-kìì-ntù</i> | / | <i>ì-biì-ntù</i> | ‘thing/s’  | class 7/8 |

In partly reduced paradigms, the reflex of *\*a* has been replaced by that of *\*i* or *\*u*. Fully reduced paradigms have the same augment vowel in all classes. De Blois (1970) cites Kimbundu H21 (*o-* in all classes), Hunde JD51 (everywhere *a-*) and Tonga M64 (always *i-*) as examples of the latter type, among many others. It is not always clear in such cases whether we are dealing with an augment according to the definition given at the beginning of this section or with a different morpheme altogether, as discussed for Eton A71, Fang A75 and Basaa A43a below. Some languages, scattered over the Bantu domain, do not have an augment, but show traces of its previous existence, which De Blois (1970: 107) calls ‘latent augments’. Typically, these traces can be found in the tone and/or vowel quality of grammatical morphemes that end in *-a*, such as the Final Vowel of verb forms, the stem *-a* of the connective relator, and the comitative-instrumental preposition *\*nà*. The Shona S10 examples in (14b, d & f) show the effect of the latent augment on the vowel of the simulative marker *sá*. There is no context in the language in which the nouns in (14a, c, e and g) appear with an initial vowel. The word for ‘tree’, for instance, is never realised as *úmùtí*. Example (14f) shows that changes in the quality of the simulative vowel cannot be ascribed to assimilation and (14h) shows that we do not find an altered simulative vowel in front of a noun of class 1a, i.e. in front of nouns that should never have had an augment.

- (12) Shona S10 (Mudzingwa & Kadenge 2013: 89)
- |    |                     |                 |
|----|---------------------|-----------------|
| a. | <i>mù-tí</i>        | ‘tree’ (cl 3)   |
| b. | <i>só = mù-tí</i>   | ‘like a tree’   |
| c. | <i>ʃî-tótà</i>      | ‘locust’ (cl 7) |
| d. | <i>sé = ʃî-tótà</i> | ‘like a locust’ |
| e. | <i>gòdô</i>         | ‘bone’ (cl 5)   |
| f. | <i>sé = gòdô</i>    | ‘like a bone’   |
| g. | <i>sèkúru</i>       | ‘uncle’ (cl 1a) |
| h. | <i>sá = sèkúru</i>  | ‘like uncle’    |

Finally, there are many Bantu languages that do not have an augment or traces of it, especially in the north-west of the Bantu area.

### 3.2 Use

The definition of the augment provided at the beginning of Section 3 is necessarily restricted to formal characteristics. The use of the augment is too language-specific to allow for a crosslinguistically valid functional definition. In many languages it depends on an intricate set

of conditions, including propositional act function (Sections 3.2.1 and 3.2.2), syntactic context (Section 3.2.3), discourse-referential properties (Section 3.2.4) and stylistics (Section 3.2.5).

In the majority of Bantu languages that have an augment, the default form of nouns is the augmented form. It is therefore usually more insightful and economic to provide conditions for its absence than for its presence (Greenberg 1978). The reason for this is that nouns are prototypically used to refer (Croft 2001: 88) and that the augment typically codes the propositional act of referring. It is therefore often absent in contexts where nouns are not used to refer and in those where nouns are inherently referential. The following is a comparative overview of the most recurrent conditions on the presence versus absence of the augment on nominal heads. These conditions are sometimes in competition and they are here presented in decreasing order of dominance.

### 3.2.1 Absence of the augment on nouns that are not used to refer

Nouns are not referring in vocatives, and also not in some cases when they are used for modification or predication. There are many augment languages where nouns used as vocatives lack an augment. In Zulu S42, for instance, “the general rule for the formation of vocatives from nouns is as follows: Elide the initial vowel of the noun-prefix” (Doke 1997: 280): *umuntu* > *muntu* (O person!).

Descriptions of individual languages are sometimes contradictory, as is the case for Ganda JE15, which has been described as lacking an augment in the vocative at least since Torrend (1891). However, vocatives do have an initial vowel augment according to Hyman & Katamba (1993), unless if followed by a second person pronoun. This may be due to dialectal differences and/or to language change.

Another context in which nouns are not used to refer is in attributive nominal predication. Predicate nouns lack the augment in Bemba M42, if they are not used to identify the subject.

(13) Bemba M42 (Givón 2001: 123)

- a. *uyu u-muu-ntu muu-puupu*  
DEM<sub>1</sub> AUG-1-person 1 -thief  
'This person is a thief.'
- b. *uyu u-muu-ntu u-muu-puupu*  
DEM<sub>1</sub> AUG-1-person AUG-1-thief  
'This person is the/a thief (I told you about).'

A reflex of the absence of the augment in this use can be found in the phenomenon of predicative lowering in Makuwa P31, a change in the tone pattern of nouns when they are used predicatively (Van der Wal 2006).

Finally, in some cases the absence of an augment could be attributed to the fact that the noun is used to modify, rather than to refer. De Blois (1970: 127) lists a number of eastern Bantu languages in which nouns lose their augment in adverbial use, i.e. when they modify a verb. In Gusii JE42, for instance, the adverb *botuko* ‘at night’ is derived from the class 14 noun *obotuko* ‘night’ by dropping the augment. Likewise, nouns tend to lack the augment when they are in the second, non-referring position of compounds, as in the Zulu S42 example in (16).

- (14) Zulu S42 (Buell 2009, citing von Staden 1973)
- |                         |              |   |                  |   |                   |
|-------------------------|--------------|---|------------------|---|-------------------|
| <i>u-m-lindimasango</i> | ‘gatekeeper’ | < | <i>u-m-lindi</i> | + | <i>a-ma-sango</i> |
| AUG-1-gatekeeper        |              |   | AUG-1-guard      |   | AUG-6-gate        |

Thus far, we saw examples of the absence of an augment where nouns are not used to refer. The flip side of the coin is that the augment can be present on word classes other than nouns, when elements that usually fulfil the propositional act of modifying are nominalised and used to refer. Examples can be found in Nande JD42 (Valinande 1984), where adjectives, possessive pronouns and other adnominal modifiers can function as NPs if they are preceded by an augment (17).

- (15) Nande JD42 (Valinande 1984: 642, 709, 714)
- a. *ǎ-mù-kirà yw-á:-yǎ* ‘his tail’ (class 9 possessor)
  - b. *ǎ-yw-á:-yǎ* ‘his one’
  - c. *ǎ-mù-tí mù-kúhí* ‘the short tree’
  - d. *ǎ-mù-kúhí* ‘the short one’

### 3.2.2 Absence of the augment on nouns that are inherently referential

Proper names tend to lack an augment for a reason that is exactly the opposite of the non-referentiality discussed in Section 3.2.1. Unlike common nouns, proper names are inherently referential. Their referential status never needs to be coded, which is also the reason why they lack a noun class prefix, i.e. why they tend to belong to the so-called class 1a. There is a small number of languages in zones G, M and S, in which the nouns of this category do have an augment, but lack a class prefix. In languages such as Rundi JD62, proper names are productively derived by omitting the augment of a common noun (Meeussen 1959, Van de Velde 2009). The nominal prefix of the source noun is morphologically integrated into the name stem. It is often pointed out that personified animals in fairy tales have the same behaviour as proper names. In fact, they are names. Another category of nouns that behave like names, but that can be distinguished from them, are kinship terms, especially the suppletive ones, which are also inherently referring in any given discourse context, due to the deictic element in their definition. Suppletive kinship terms such as *kitàáwé* ‘his father’ and *bázé* ‘my husband’ never take an initial vowel augment in Ganda JE15 (Hyman & Katamba 1993: 222). When borrowings are integrated into class 1a on formal grounds, i.e. due to their lack of a recognisable class marker, they are often fully assimilated to proper names, also in their incompatibility with the augment.

### 3.2.3 Absence of the augment in certain syntactic environments

Together with the propositional act function, the syntactic context is the most widespread conditioning factor for the presence or absence of the augment. Usually, it determines whether the use of the augment is prohibited, obligatory or syntactically optional, but only for those nouns that are not inherently augmentless for reasons specified in Sections 3.2.1 and 3.2.2. Hence, the propositional act function generally outranks the syntactic context as a conditioning factor. On clause level, the augment is often absent or contrastive, i.e. syntactically optional, on the object of a negative verb form. In Ganda JE15, the augment is likewise absent on an object noun when it is under focus as in (18b) (Hyman & Katamba 1993).

- (16) Ganda JE15 (Hyman & Katamba 1993: 228)
- a. *y-à-gúl-ìr-à*                      *à-bá-àná*                      *è-bí-tábó*  
 SP<sub>1</sub>-PST-buy-APPL-FV    AUG-2-child    AUG-8-book  
 ‘He bought the children books.’
- b. *y-à-gúl-ìr-à*                      *bá-àná*                      *è-bí-tábó*  
 SP<sub>1</sub>-PST-buy-APPL-FV    2-child                      AUG-8-book  
 ‘He bought THE CHILDREN books.’

On the level of the noun phrase, head nouns often lose their augment in the presence of (certain types of) preposed modifiers. The augment can then either function as a phrasal affix and appear on the prenominal modifier (19) or be left out (20).

- (17) Nande JD42 (Valinande 1984: 819)
- a. *ǎ-mù-tí*                      ‘the tree’  
 b. *ǎ-yù-ntì mù-tí*                      ‘the other tree’

- (18) Zulu S42 (Buell 2009)
- a. *i-ndawo leyo*                      ‘that place’  
 b. *le ndawo*                      ‘this place’

Locative prefixes are also often incompatible with the augment. In many languages they are not preceded by the augment of their own locative class, nor are they followed by the augment of the noun to which they attach, but exceptions to both types of absence of the augment occur (De Blois 1970: 117-119, Grégoire 1975: 156-169).

A few postnominal modifiers have been found to be incompatible with the augment too, at least in parts of the Bantu domain, such as the interrogative ‘how many’ (Grégoire & Janssens 1999: 417).

A completely different type of syntactic conditioning within the noun phrase is attested in Basaa A43a and in the A70 languages, where the non-augmented form is the default form of nouns and an augment appears if and only if the noun is restrictively modified by certain adnominal modifiers. The augmented form can therefore be analysed as a construct form of the noun (Van de Velde 2017). The set of modifiers that require the augmented/construct form is language specific, but always includes postnominal demonstratives. Note, however, that the form of the augment in these languages is an invariant vowel *í-* or *é-*, sometimes reduced to a floating high tone, which makes it hard to determine whether this form is cognate with the forms identified as the augment in other Bantu languages.

### 3.2.4 Absence of the augment with indefinite or non-specific nouns

In many Bantu languages, an augmented form of the noun can contrast with a non-augmented form, at least in some syntactic contexts, sometimes including nouns in isolation. The augment is usually described as expressing definiteness or specificity in such cases. Unfortunately, evidence for this tends to be restricted to isolated examples accompanied by a translation equivalent in French or English that contains a definite article for the Bantu form with an augment and an indefinite article for that without an augment (21). An oft-cited example of a language in which the augment is claimed to express definiteness is Zamba C322 (Bokamba 1971, Kamanda Kola 1994). Other examples include Ngazidja G44a (cf. Patin et al, this volume) and Nande JD42 (Valinande 1984: 432). These three languages have

in common that the use of the augment on nouns modified by a postnominal demonstrative is obligatory.

(19) Zamba C322 (Kamanda Kola 1994: 402)

- a. *bá-bà-áná*    *bà-nd'*    *ó-dǎn-á*  
 AUG<sub>2</sub>-2-child    SP<sub>2</sub>-PROG    INF-play-FV  
 'The children are playing.'
- b. *bà-áná*    *bà-nd'*    *ó-dǎn-á*  
 2-child    SP<sub>2</sub>-PROG    INF-play-FV  
 'Children are playing.'

Patin (2017) does provide evidence for the use of the augment as a marker of definiteness in Shingazidja, in contexts where its use is syntactically optional. Similar evidence may exist for languages such as Zamba and Nande. Note that in none of these languages the augment is entirely dedicated to the expression of definiteness. In Zamba C322, for instance, it also functions as a nominaliser of modifiers and as an intensifier on nouns (including infinitives).

In the absence of clear evidence, some scepticism is warranted by the fact that relying on translation equivalents in European languages in order to establish whether the augment expresses definiteness can be treacherous. An example of this can be found in Fang Ntumu A75A, where the augment has been claimed to be a marker of definiteness (Ondo Mebiame 2001). However, Fang being an A70 language, the conditioning for the appearance of the augment is basically syntactic (see my claim in Section 3.2.3). The augment is triggered by the presence of certain types of modifiers that can be functionally characterised as anchoring (or localising) (Rijkhoff 2002: 173-212).<sup>iii</sup>

Anchoring modifiers are those modifiers that allow the hearer to identify the referent of the noun phrase in the world of discourse by locating it in space or by linking it to an already identified entity. Modifiers that allow both an anchoring and a non-anchoring use are responsible for the impression that the augment marks definiteness, since the presence of an anchoring modifier changes the discourse-referential properties of a nominal constituent. The only concrete indication for calling the augment a definiteness marker in Fang Ntumu A75A is that the modifier *-fǎ* translates as 'the other' when the head noun takes an augment and as 'another' in the absence of an augment. However, it is clear that the presence of the augment changes the type of modification from non-selective to selective (and therefore anchoring), which merely implies a definite interpretation of the head noun. This analysis is confirmed by word order in the nominal constituent, in that anchoring modifiers obligatorily follow other types of modifiers in the A70 languages. Adnominal *-fǎ* in Fang Ntumu A75A can precede or follow a cardinal number in the absence of an augment (22a-b), but it has to follow when the head noun is augmented (22c-d) (Van de Velde 2017).

(20) Fang Ntumu A75A (Van de Velde 2017)

- a. *mətá mǎ'béŋ mǎ'fǎ*  
*mə-tá*    *mǎ-běŋ*    *mə-fǎ*  
 6-pile    PP<sub>6</sub>-two    PP<sub>6</sub>-other  
 'Two other piles.'
- b. *mətá mǎ'fǎ mǎ'béŋ*  
*mə-tá*    *mə-fǎ*    *mǎ-běŋ*  
 6-pile    PP<sub>6</sub>-other    PP<sub>6</sub>-two
- c. *mətá mǎ'béŋ mǎ'fǎ*



- <sup>H</sup>-*mə-tá*      *mə-běŋ*      *mə-fě*  
 AUG-6-pile    PP<sub>6</sub>-two    PP<sub>6</sub>-other  
 ‘The two other piles.’  
 d. \**mətá məfě məběŋ*

The role of the augment as a marker of specificity (or referentiality) in languages in which augmentless nouns can contrast with augmented nouns has been shown more convincingly, most notably for Bemba M42 (Givón 2001: 453). Nouns that are in the scope of a realis proposition are necessarily specific/referring, and therefore must take the augment in Bemba.

(21) Bemba M42 (Givón 2001)

- a. *a-a-som-ene*                      *i-ci-tabo*  
 SP<sub>1</sub>-PST-read-PFV              AUG-7-book  
 ‘She read a/the book.’  
 b. \**a-a-som-ene ci-tabo*  
 \*‘She read a book (not a specific one).’

In other contexts, where indefinite nouns are not necessarily referring, the absence of the augment signals non-referentiality.

(22) Bemba M42 (Givón 2001)

- a. *a-a-fwaay-ile*              *u-ku-soma*              *i-ci-tabo*  
 SP<sub>1</sub>-PST-want-PFV    AUG-15-read              AUG-7-book  
 ‘She wanted to read a/the book’ (a specific book)  
 b. *a-a-fwaay-ile*              *uku-soma*              *ci-tabo*  
 SP<sub>1</sub>-PST-want-PFV    AUG-15-read              7-book  
 ‘She wanted a book to read (any book).’  
 c. *ta-a-a-som-ene*                      *i-ci-tabo*  
 NEG-SP<sub>1</sub>-PST-read-PFV              AUG-7-book  
 ‘He didn’t read the book.’  
 \*‘He didn’t read a book.’  
 d. *ta-a-a-somene*                      *ci-tabo*  
 NEG-SP<sub>1</sub>-PST-read-PFV              7-book  
 ‘He didn’t read any book.’

The absence of the augment to signal non-referentiality in these examples is related to the cases of lack of an augment discussed in Sections 3.2.1 and 3.2.2. Moreover, the syntactic conditionings on clause level discussed in Section 3.2.3 must result from differences in frequency between a referential and a non-referential reading in contexts where languages such as Bemba M42 allow the presence of the augment to contrast with its absence.

### 3.2.5 Stylistic and prosodic conditionings

The use of the augment has been reported to be subject to stylistic conditions in some languages. Thus, according to Carter (1963, cited via Greenberg 1978: 254), speakers of Tonga M64 find the excessive use of the augment undignified, and there are considerable differences in the frequency of augment use according to text style. Likewise, a comparative study of eight languages of zone G in which the augment is on its way out or has lost its traditional type of distribution, has shown that the use of the augment depends on register and

is sometimes used by communities to differentiate their speech from that of neighbouring groups (Aunio & Petzell 2013).

Finally, Bantu languages tend to “care about” prosodic well formedness of words, affixes and stems, and there are numerous examples of the presence of an augment solely motivated by a necessity to comply to minimality constraint on words, for instance. Thus, Gisu JE31a nouns have an augment in the classes where the pronominal prefix differs segmentally from the nominal prefix, i.e. classes 3, 4, 6 and 10 (see Section 3.1 for this prosodic conditioning). In the other classes, an augment appears only before monosyllabic or vowel-initial stems (De Blois 1970: 94), i.e. *u-mu-ndu* / *ba-ba-ndu* ‘person / people’ (cl. 1/2) vs. *ba-kana* ‘girls’ (cl. 2).

### 3.3 Origin

It is undeniable that augments or traces of them can be found throughout the Bantu area, also in the north-west (Grégoire & Janssens 1999). Meeussen (1967: 99) tentatively reconstructs it as a pronominal demonstrative form, “a separate word, identical in form with the pronominal prefix, and used as a weak demonstrative, or rather anaphoric, in affirmative, non-predicative constructions, with definite meaning.” Greenberg (1978: 254) characterises the Proto-Bantu augment as (probably) a “stage II” article. Stage II refers to a phase in a continuous evolution of markers that start out as demonstratives (stage zero) and become definite articles, i.e. pragmatically conditioned markers of definiteness (stage I). The presence of these markers subsequently becomes grammatically conditioned. They tend to become obligatory with common nouns, except in a number of language-specific grammatical contexts (stage II). The fact that the article has lost its functional load at stage II can either lead to its generalisation, so that it becomes a mere nominal marker (stage III), or to its erosion and disappearance. Thus, both Meeussen (1967) and Greenberg (1978) reconstruct the Proto-Bantu augment at a rather advanced stage of grammaticalisation, assuming that syntactic restrictions on its distribution already existed before the Bantu languages dispersed. However, it is very difficult to determine whether Proto-Bantu had an augment and, if it did, at what stage of the Greenberg cycle it has to be situated. The reason is that there is ample evidence for the existence of multiple cycles of augment creation and loss in Bantu and beyond. Some of this evidence is presented in the remainder of this section.

For instance, we saw that the augment in Zamba C322 is claimed to express definiteness, i.e. that at least in some of its uses it functions as a Stage I article, which puts it earlier in the cycle of grammaticalisation than the proposed reconstructed augment. Moreover, Zamba is the only language in its region and genealogical subgroup that has a fully functional augment (Kamanda Kola 1994). The other languages of zone C, such as Doko C301 (Twilingiyimana 1984), show at most some traces. The Zamba C322 augment should be interpreted as an innovation, rather than an exceptional retention. Formal evidence for this is provided by the tone of the nominal prefix in Zamba, which Kamanda Kola (1993) analyses as underlyingly high. Since nominal prefixes are low in Proto-Bantu and the contemporary languages surrounding Zamba and since the augment has a high tone, the most straightforward explanation for the high tone on Zamba nominal prefixes is that it is the reflex of an augment that generalised into a Stage III article. Consequently, the current segmental augment must be a renewal. There are other Bantu languages in which the basic tone pattern of nouns includes a generalised historical augment High. An interesting example is Yombi H16, where a tonal

distinction between definite and indefinite nouns has been innovated by the creation of a new indefinite form, derived from the form of nominal predicates (Blanchon 1998).

Indications for the renewal of augments can also be found in languages with two paradigms of augments, one of which typically has a CV-shape and the other a V- shape. This can be found in Safwa M25, Kinga G65 and Nyakyusa M31, where the CV-augment is claimed to have a special “emphatic” use (De Blois 1970: 98, confirmed for Nyakyusa by Bastian Persohn pers. comm.)

(23) Nyakyusa M31 (Bastian Persohn, pers. comm.)

<i>ʊ-mu-ndʊ</i>	‘the person’
<i>ʊ-mu-ndʊ ʊ-ju</i>	‘this person’
<i>ju-mu-ndʊ</i>	‘the very person’

The Zamba C322 augment too can be used as an intensifier (26b) and preposed demonstratives have been characterised as “emphatic” in Nkore JE13, Kanyok L32, Bemba M42 and Bolia C35b, as opposed to their postposed counterparts (Van de Velde 2005). In fact, it is not always straightforward to know whether we are dealing with an augment or rather with a pronominal demonstrative.

(24) Zamba C322 (Kamanda Kola 1994: 403)

- a. *ândó wà Mádángá*  
‘He is from (the village of) Madanga.’
- b. *ândó wà Mádángá (má-mádángá)*  
‘He is from Madanga itself, from the very village of Madanga.’

The existence of augment-like morphemes in Wide Bantu or Bantoid languages has been adduced as evidence for the antiquity of the augment in Bantu, but it can just as well show that Niger-Congo languages with elaborate noun class systems are very prone to develop augment-like categories and that this strong tendency constantly feeds cycles of grammaticalisation within Narrow Bantu as well. To give only one example, the Atlantic language Keerak has a morpheme that Segerer (2015: 107) calls ‘minimal determiner’. It is an enclitic =*aC* (where C is identical to the consonant of the noun class prefix), which appears in most syntactic contexts, including in isolation. In the few contexts where the minimal determiner is optional, e.g. before an adjective, its presence may express definiteness.

#### 4 ADNOMINAL MODIFIERS

Adnominal modifiers normally agree in noun class with their head noun. They are traditionally classified according to their paradigm of agreement markers. Meeussen (1967: 97) reconstructed the five paradigms provided in Table 1: nominal prefixes, enumerative prefixes, pronominal prefixes, subject prefixes and object prefixes, the first four of which are used to mark agreement within the noun phrase. Contemporary languages may distinguish either more or less different paradigms than those reconstructed in Proto-Bantu. This section provides a brief overview of some types of adnominal modifiers.

## 4.1 Demonstratives

Bantu languages tend to be rich in demonstratives, both in types and in frequency of use. They typically have three or four series of demonstratives, which, in their exophoric use, express different degrees of spatial distance from the speaker and/or the hearer.

(25) Nande JD42 (Valinande 1984: 787-792)

- a. *ǎ-mù-sìkàà ǎ-nǎ* ‘this girl (within reach of the speaker)’
- b. *ǎ-mù-sìkàà ǎ-lià* ‘that girl (far from speaker and hearer)’
- c. *ǎ-mù-sìkàà ò-iù* ‘this girl (close to the speaker, but out of their reach)’
- d. *ǎ-mù-sìkàà ǎ-iù-ǎ* ‘that girl (close to the hearer)’

Formally, demonstratives minimally contain the pronominal prefix, sometimes accompanied by a stem and a full or partial repetition of the agreement marker. Reconstructed demonstrative stems include *\*-nóò* [close to speaker] and *\*-díà* [far from speaker and hearer] (Meeussen 1967: 107). Recent research on demonstratives in Bantu, most notably by Nicolle (2012, 2014), has focused on their discourse functions, an important aspect of their use that is hardly mentioned in grammatical descriptions. In a comparative corpus study of ten Eastern Bantu languages, Nicolle (2014) identified four different discourse level functions of demonstratives, viz. indicating the so-called activation status of participants (distal demonstratives reactivate participants after a period of inactivity in languages such as Jita JE25); indicating agentivity (agentive participants are modified by distal demonstratives, others by referential demonstratives); “text structuring” (distinguishing between different types of participants, delimiting episodes) and thematic development (indicating important developments in the narrative).

## 4.2 Connectives (genitives)

Nominal possessors are introduced by a relator that typically consists of a pronominal prefix and a stem *-a*, the tone of which is identical to that of the prefix. Bantuists call this relator the ‘connective’, sometimes also ‘connexive’ or ‘associative’ marker.

(26) Nyamwezi F22 (Maganga & Schadeberg 1992: 89)

- m̄-zuna*                      *w-aa-m̄-kíuma*  
 1-younger\_sister    PP<sub>1</sub>-CON-1-woman  
 ‘the younger sister of the woman’

Van de Velde (2013) provides a typology of connective constructions in the Bantu family, which shows, among other things, that the expression of linguistic possession is only one of their multiple functions. In fact, possessive examples such as (28) may be relatively rare in discourse in Northwestern Bantu languages, since many of them prefer external possessor constructions, in which the possessor is expressed as an argument of the verb. In Orungu B11b, for instance, there are almost no restrictions on external possessors and it is not unusual to find multiple external possessors in one clause (29b).

(27) Orungu B11b (Odette Ambouroué p.c.)

- a. *áfè wádyónì ìgúgè ñ ínàgò yámì*  
*á-fè*                      *wá-á-dyón-ì*                      *ì-gúgè*                      *ñ-á*                      *ì-nàgò*                      *ìy-ámì*  
 2-burglar.DTP    SP<sub>2</sub>-PRF-break-PRF    5-door.DTP    PP<sub>5</sub>-CON    AU<sub>9</sub>-9.house.DTP    PP<sub>9</sub>-POSS<sub>1SG</sub>

‘The burglars have broken down the door of my house.’

b. *áfè wádyónì myé nág ìgúgè*

*á-fè wá-á-dyón-ì myé nágò ì-gúge*

2-burglar.DTP SP<sub>2</sub>-PRF-break-PRF 1SG 9.house.DTP 5-door.DTP

‘The burglars have broken down the door of my house.’ (lit. ‘have broken me the house the door’)

### 4.3 Possessive pronouns

Possessive pronouns consist of a stem and a pronominal prefix that agrees with the possessee. In most Bantu languages, the form of the stem depends on the person and number of the possessor, but in some languages it agrees with the possessor in noun class as well, giving rise to huge paradigms. These languages can be found in different regions, e.g. Herero R31 (Möhlig *et al.* 2002: 60), Binza C321 (Van Leynseele 1977: 35) and Ha JD66 (Harjula 2004: 69) and Nande JD42 (Valinande 1984). Possessive pronouns for non-human possessors are connective constructions with a pronominal possessor.

(28) Herero R31 (Möhlig *et al.* 2002: 60)

a. *ò-mù-tí n-ò-ví-yàò vy-á-‘w-ó*  
AUG-3-tree and-AUG-8-leave PP<sub>8</sub>-CON-PP<sub>3</sub>-PPR  
‘The tree and its leaves.’

b. *ò-tjì-kúnìnò n-ò-mí-tí vy-á-‘ty-ó*  
AUG-7-garden and-AUG-4-tree PP<sub>4</sub>-CON-PP<sub>7</sub>-PPR  
‘The garden and its trees.’

The Bantu languages typically have a number of suppletive kinship terms, in which a kinship term and a possessive pronoun are merged. A comparative-historical study of the form of possessive pronouns can be found in Kamba-Muzenga (2003).

### 4.4 Numbers

Cardinal numbers from ‘one’ to ‘five’ usually agree with the noun they modify, sometimes using a dedicated set of numeral agreement markers. From ‘six’ upwards they tend to be uninflected (Stappers 1965, Meeussen 1967). Ordinal numbers are usually expressed by means of a connective construction with a cardinal number in the modifier position (Polak-Bynon 1965).

(29) Makwe G402 (Devos 2008: 133)

*sìkù y-a-tàànò*  
9.day PP<sub>9</sub>-CON-five  
‘the fifth day’

### 4.5 Relative clauses

Some parameters of variation in the structure of relative clauses across Bantu languages are briefly discussed by **Downing & Marten (this volume)**. Two additional parameters of variation related to the agreement prefix of the relative verb are especially relevant for this section: its paradigm and its controller (Nsuka-Nkutsi 1982). Relative verbs can take an agreement prefix of the verbal (subject) paradigm or of the pronominal one. In non-subject

relatives, the controller of agreement can be either the subject of the relative verb or the relativised noun. There is a partial correlation between these two parameters, in that relative verbs that always agree with their subject, never take a pronominal prefix. The reverse correlation is strong, but not without exceptions: the great majority of relative verbs that agree with the relativised noun take a pronominal prefix rather than a verbal prefix. A straightforward explanation for this correlation is that the pronominal prefix in relative verbs originates in a demonstrative that was used to nominalise verb forms.

#### 4.6 Adjectives

Bantu adjectives are defined as agreeing words that take an adjectival prefix, a paradigm that is fully or nearly identical to that of the nominal prefixes. The contents of this Bantu-specific word category differs from language to language. In some languages, such as Ngazidja G44a (cf. Patin et al. this volume), it also contains numbers. Most Bantu languages have a limited set of adjectives (e.g. eight in Rwanda DJ 62), some have none at all (e.g. Eton A71, Mongo C61). A small number of qualifying adjectives can be reconstructed into Proto-Bantu: *\*-bû* ‘bad’, *\*-bîci* ‘unripe’, *\*-dâi* ‘long’, *\*-dîtò* ‘heavy’, *\*-nénè* ‘big’, *\*-pâi* ‘new’ and *\*-tádí* ‘long’ (Meeussen 1967, Baka 2000, Bastin et al. 2002, Baka 2005). In languages with an open class of adjectives, often most are derived from verbs, usually change-of-state verbs. The most common productive deverbal derivation types involve the suffixes *\*-é* (Bastin 1989) and *\*-ú* (Schadeberg 2003: 81).

The scarcity of adjectives in the majority of languages is compensated for by alternative strategies for nominal qualification. A common strategy is to use a connective construction, in which the modifying noun can be a property denoting (32) or entity denoting (33) noun or the infinitive of a change-of-state verb (34).

- (30) Ha JD66 (Harjula 2004: 135)  
*umutaama wikigongwe*  
*u-mu-taama u-a i-ki-gongwe*  
 AUG-1-old.man PP<sub>1</sub>-CON AUG-7-kindness  
 ‘A sympathetic old man.’

- (31) Mongo C61 (Hulstaert 1966: 247)  
*e-kútu ě-a = n-dɔsɔ*  
 7-calabash PP<sub>7</sub>-CON=10-pore  
 ‘A porous calabash.’

- (32) Digo E73 (Nicolle 2013)  
*ng’ombe z-a ku-ond-a*  
 10.cow PP<sub>10</sub>-CON INF-become.thin-FV  
 ‘Thin cows.’

In some north-western Bantu languages, the qualifying noun is construed as the morphosyntactic head of the connective construction, in a construction type that instantiates possessee-like qualifier constructions. Such constructions are typical in a large area of northern Sub-Saharan Africa.

- (33) Eton A71 (Van de Velde 2008: 214)

*ìvèvèz m̀pég í'té kù*  
*ì-vèvèz H=N-pég í-Lté L-kù*  
 7-light PP7.CON=3-bag SP7-PRS INF-fall  
 'The light bag falls.'

Some Southern Bantu languages (S30-40) have a word class traditionally called 'relatives', which has the functions and distributional potential of adjectives. These words have a stem preceded by an agreement prefix of the paradigm of subject markers on verbs. As most other adnominal modifiers in these languages, they are obligatorily preceded by a linker in attributive use. Creissels (2014), who prefers the term 'new adjectives', argues that they emerged as a new word class from the use of nouns as descriptive predicates. Example (36) contrasts an old type adjective (36a) with a new adjective (36b) in Tswana S31.

- (34) Tswana S31 (Creissels 2014)
- a. *m̀-śádì jó m̀-léèlé*  
 1-woman LNK<sub>1</sub> NP<sub>1</sub>-tall  
 'A tall woman.'
- b. *m̀-śádì jó ú-bùtʰálí*  
 1-woman LNK<sub>1</sub> SP<sub>1</sub>-clever  
 'A clever woman.'

Other strategies for adnominal qualification include the use of relative clauses and participles.

## 5 THE STRUCTURE OF THE NOUN PHRASE

The Bantu languages show a lot of variation between them in the word order patterns of the noun phrase. Some of the patterns that can be found in Bantu, such as N DEM ADJ NUM or N NUM DEM ADJ, are typologically very unusual or have never been attested in the typological literature (see for instance Greenberg 1963, Hawkins 1983, Rijkhoff 2002, Cinque 2005) and (Wide) Bantu languages have had a heavy impact on the way universals of word order in the NP have been formulated. Thus, Greenberg's universal 20 states that demonstratives, numbers and adjectives occur in this order (i.e. DEM NUM ADJ) when they all precede and in the same or in the exact opposite order if they all follow (Greenberg 1963). The fact that Greenberg allows the order DEM NUM ADJ among postnominal modifiers is apparently due to his knowledge of Gikuyu E51, which is not part of his 30-language sample. Hawkins (1983) later reformulated Universal 20, based on a sample of more than 300 languages, stating that nothing can be predicted about the mutual ordering of postnominal demonstratives, numbers and adjectives. The Wide Bantu languages Aghem and Noni, which allow N ADJ DEM NUM and N DEM NUM ADJ respectively, are responsible for this further weakening of Greenberg's universal 20.

In light of this prominence of the Bantu languages in the general literature, it is somewhat surprising that there exists no comparative study of the structure of the noun phrase in Bantu. The following subsections are a first attempt. Comparing the word order in the NP of individual Bantu languages to typological findings is complicated by three factors. First, the three adnominal modifiers to which typologists have restricted their studies exclude the possessive pronoun, the position of which is noteworthy in Bantu. Second, it is not a priori clear which kind of definition for the different types of adnominal modifiers would produce the most insightful results in a comparative study, a functional one or a formal one. For

instance, according to Rijkhoff's (2002) (functional) layered model of the NP, we expect numbers to be serialised differently from qualifiers, whatever the language-specific grammatical category to which they belong. In many Bantu languages, numbers from 'one' to 'five' are analysed as Adjectives, since they take the same paradigm of agreement markers as qualifying adjectives. The two are therefore often not distinguished in schematic representations of word order in the NP. The third complicating factor in comparing word order in Bantu NPs to what is found in the typological literature is that many Bantu languages allow for considerable syntactic freedom within the NP.

## 5.1 Common word order patterns

Most adnominal modifiers are postnominal, but demonstratives (37a), various quantifiers (37b) and/or the modifier 'other' (37c) are prenominal in many languages. This is also sometimes the case for possessive pronouns, for which prenominal position is usually (or always?) syntactically optional and pragmatically marked, e.g. in Songye L23 (Stappers 1964) and Makaa (38b).

(35) Ha JD66 (Harjula 2004: 75, 131)

- a. *izo súka zi-bíri*  
DEM<sub>10</sub> 10.hoe PP<sub>10</sub>-two  
'Those two hoes.'
- b. *burú mu-ntu*  
each 1-person  
'Each person.'
- c. *u-wú-ndi mú-si*  
AUG-PP<sub>3</sub>-other 3-day  
'Another day'.

(36) Makaa A83 (Heath 1998: 3)

- a. *me angane dug boog j-am*  
1SG PROG.NEG see 7.hoe PP<sub>7</sub>-POSS<sub>1SG</sub>  
'I don't see my hoe.'
- b. *me ke gule ne j-am boog, wo ke ne gwoo boog*  
1SG go hoe with PP<sub>7</sub>-POSS<sub>1SG</sub> 7.hoe 2SG go with PP<sub>7</sub>.POSS<sub>2SG</sub> 7.hoe  
'I am going to hoe with MY hoe, you go with YOUR hoe.'

Among postnominal modifiers, the position of the possessive pronoun is remarkable, as in many languages it is obligatorily adjacent to the noun, i.e. preceding any other postnominal modifiers. Again, this is typologically highly unusual, since possessive pronouns are typically anchoring modifiers, which are nearly universally placed further away from the head noun than classifying, qualifying and quantifying modifiers (Rijkhoff 2008). The same is true for postposed demonstratives, which in some languages are serialised immediately after the noun, or after the possessive pronoun, preceding other modifiers. Some examples of the order of demonstratives, possessive pronouns, numbers and adjectives are: Digo E73 DEM N POSS NUM ADJ (Nicolle 2013), Nande JD42 N POSS NUM ADJ DEM (Valinande 1984: 633), Nkore-Kiga JE13-14 N POSS DEM ADJ NUM (Taylor 1985: 55 who adds that "this order is not rigidly adhered to"), and Orungu B11b N POSS ADJ DEM (Odette Ambouroué pers. comm., see below for the position of NUM).



The preference for the immediate postnominal position of possessive pronouns is so strong in some Bantu languages, that it gives rise to external possession within the noun phrase, as in the Mbugwe F34 example in (39) (cf. [Wilhelmson, this volume](#)). Semantically, the possessive pronoun modifies the noun ‘body’ in the connective construction, but formally it follows the head noun of the NP, with which it also agrees.

- (37) Mbugwe F34 ([Wilhelmson, this volume](#))  
*n-yèèngò jì-ááné jì-óónsè jì-á mò-vèrè*  
 10-joints PP10-POSS1SG PP10-all PP10-CON 3-body  
 ‘All the joints of my body.’

The core position that the possessive pronoun occupies in the NP structure of many Bantu languages, is also reflected in its agreement properties. As was said in Section 2.2.2.2, possessive pronouns are at the top of the agreement hierarchy in some Bantu languages, since they are the only agreement targets that can have syntactic agreement with animate controllers. Strangely, syntactic agreement is restricted to certain noun classes. In Swahili G42, human controllers trigger animate agreement on all agreement targets, except on the possessive pronoun, if the head noun has a class 9/10 prefix.

- (38) Swahili G42 (Van de Velde 2006b: 195)  
*N-dugu y-angu a-me-anguk-a*  
 9-brother PP9-POSS1SG SP1-PFV-fall-FV  
 ‘My brother has fallen.’

Likewise, in Fe’Fe’ (Wide Bantu), possessive pronouns are the only agreement target for which the full range of noun classes are differentiated (Hyman 1972).

## 5.2 Word order variation in the NP

In many Bantu languages, the mutual ordering of some or all postnominal modifiers is syntactically free, represented by curly brackets around the interchangeable modifiers in the schemes that follow: Basaa A43a N {POSS, NUM, ADJ} DEM (DEM and POSS can both or either be prenominal as well) (Hyman 2003); Ha JD66 DEM QUANT N {POSS, ADJ, NUM} {REL, PART} (Harjula 2004); Eton A71 N {POSS, NUM, CON} REL DEM (Van de Velde 2008: 227). The position of cardinal numbers is syntactically free in Orungu B11b, and in their presence the otherwise rigid order of DEM and ADJ becomes free as well. The possessive pronoun remains in immediate postnominal position (Odette Ambouroue, pers. comm.): N POSS {NUM, ADJ, DEM}.

The possible semantic and information structural implications of alternative word orders in the noun phrase have received little attention in the literature. For Rwanda JD61, Wilkins & Kimenyi (1975) claim that the order of adnominal modifiers is determined by a generality hierarchy, such that the modifiers that contribute most to the identification of the NP’s referent are placed at the back. This hierarchy can explain both the default order in Rwanda and departures from it. In the following example, the relative clause can optionally precede the adjective only if it expresses a habitual activity, i.e. a more or less permanent quality.

- (39) Rwanda JD61 (Wilkins & Kimenyi 1975: 160)  
 a. *u-mu-gabo mu-re-mu-re u-ririimb-a n’ u-mu-txwaare*

- AUG-1-man NP<sub>1</sub>-tall-NP<sub>1</sub>-tall PP<sub>1</sub>-sing-FV is AUG-1-chief  
 ‘The tall man who sings is the chief.’
- b. *u-mu-gabo uririimba mu-re-mu-re n’ u-mu-txwaare*  
 AUG-1-man PP<sub>1</sub>-sing-FV NP<sub>1</sub>-tall-NP<sub>1</sub>-tall is AUG-1-chief  
 ‘The TALL man who sings, is the chief.’

### 5.3 Towards a historical explanation for the uncommon word order patterns in Bantu NPs

A possible explanation for the typologically unusual word order patterns in Bantu is that in many languages the noun phrase is or was not fully integral, and that the relation between a head noun and some of its semantic modifiers is or was appositional. There are several types of indications for the validity of this hypothesis. First, as we saw in Section 3.2.1, the augment is often used to nominalise modifiers. The fact that several types of adnominal modifiers can take the augment in some Bantu languages is an indication that these modifiers are or were nominalised and in an appositional relation with the head noun. In Bemba M42, for instance, modifiers are non-restrictive – and therefore arguably in loose apposition – when they take an augment, and restrictive when they don’t (Givón 1974: 132, Kasonde 2009: 167).

(40) Bemba M42 (Givón 1974: 132, Kasonde 2009: 167)

- a. *a-ba-ntu ba-suma*  
 AUG-2-person NP<sub>2</sub>-good  
 ‘The good people.’
- b. *a-ba-ntu a-ba-suma*  
 AUG-2-person AUG-NP<sub>2</sub>-good  
 ‘The people, who were (all) good’; ‘The people, the good ones.’
- c. *a-ba-ana ba-andi*  
 AUG-2-child PP<sub>2</sub>-POSS<sub>1SG</sub>  
 ‘My children.’
- d. *a-ba-ana a-ba-andi*  
 AUG-2-child AUG-PP<sub>2</sub>-POSS<sub>1SG</sub>  
 ‘The children, those that are mine.’

A number can only be inserted between the noun and the adjective if the latter is augmented/apposed (Kasonde 2009).

(41) Bemba M42 (Kasonde 2009)

- a. *à-báá-ntù bà-bilì á-bà-kúlú*  
 AUG-2-person NP<sub>2</sub>-two AUG-NP<sub>2</sub>-big  
 ‘The two men, the big ones.’
- b. *\*à-báá-ntù bà-bilì bà-kúlú*

The use of appositive structures involving a nominalised modifier is not necessarily reserved for non-restrictive modification. It can also distinguish the anchoring use of modifiers from other uses. In Babole C101, for instance, the connective relator has a short and a long form (Leitch 2003). The long form consists of the short form, prefixed by an augment. It is used when the connective construction expresses linguistic possession, i.e. when it has an anchoring function. Connective modifiers with a classifying or qualifying function take the

short, non-augmented form of the connective relator. The same distinction exists in Zamba C322 (Kamanda Kola 1994).<sup>iv</sup>

(42) Zamba C322 (Kamanda Kola 1994: 405)

- a. *ímúntòdù mwà mwàsì* ‘the navel of a woman’
- b. *ímúntòdù ímwà mwàsì* ‘the woman’s navel’

The creation of such appositive structures may be instigated by the tendency, noted for Rwanda JD61 in Section 5.2, to put the modifiers that contribute most to the identification of the NP’s referent at the right edge of the phrase.

Second, there is morphosyntactic evidence for analysing certain modifiers as extraphrasal. In Ganda JE15, for instance, an adnominal modifier can be “extraciated.” It then follows any other modifiers and takes the augment, also if the noun phrase from which it is extracted follows a negative verb form, which does not allow for the augment (Hyman & Katamba 1993).

(43) Ganda JE15 (Hyman & Katamba 1993)

- a. *tè-y-à-láb-à*                    *bì-tábó* *bì-néné* *bì-nó*  
NEG-SP<sub>1</sub>-PST-see-FV 8-book NP<sub>8</sub>-big PP<sub>8</sub>-DEM  
‘He didn’t see these big books.’
- b. *tè-y-à-láb-à*                    *bì-tábó* *bì-nò* *è-bì-néné*  
NEG-SP<sub>1</sub>-PST-see-FV 8-book PP<sub>8</sub>-DEM AUG-NP<sub>8</sub>-big  
‘He didn’t see these big books.’

In Nen A44, which has verb-final clausal syntax, part of a discontinuous object NP can even end up after the verb, marking contrastive focus on the extracted modifier (Mous 2003).

(44) Nen A44 (Mous 2003: 345)

- mè-ná*                    *ìmità*                    *yè*                    *m<sup>w</sup>ànífí*                    *índí*                    *mè-ńén*                    *ò*                    *hè-lóbátò*  
1SG-HOD.PST 9.calabash CON<sub>9</sub> 6.water give.HOD.PST NP<sub>9</sub>-big LOC 19.child  
‘I gave the BIG water calabash to the child.’

Recall that the augment is a renewal of agreement morphology, that cycles of augment creation and disappearance are likely to have occurred and that eventually all agreement markers within the noun phrase must have started as demonstratives used to nominalise adnominal modifiers.<sup>v</sup> Consequently, the structure of the noun phrase in contemporary languages can be the result of cycles of appositional extraposition of modifiers followed by subsequent reintegration, giving rise to multiple layers of modifiers around a nuclear NP. This scenario should be tested in a thorough comparative study, which should pay special attention to prosodic clues for the syntactically layered nature of Bantu NPs. In loose apposition, an anchor is typically separated from the appositive by means of a prosodic boundary, as O’Connor & Patin (2015) have demonstrated for Ngazidja G44a. Differences between types of modifiers in prosodic bonding with their head in the Bantu languages could be reflexes of such boundaries. In Binza C321, for instance, there is high tone dissimilation with the last syllable of a nuclear NP (N (ADJ)) on a following demonstrative or possessive pronoun, but not on a following quantifier or connective (Van Leynseele 1977). In Makwe G402, adnominal modifiers can be classified into conjoint, disjoint and conjoint/disjoint, depending on whether they must, cannot or may form a phonological phrase with the immediately preceding head of the NP (Devos 2008: 377-382). Finally, nouns that end in two high tones

become high-low before a pause in Tswana S31. They acquire the same final tone pattern when preceding certain types of adnominal modifiers. Since the final HL pattern is phonologically unpredictable in these cases, Creissels (2009: 79) rightly analyses it as a construct form of the noun.

## REFERENCES

- Aunio, L. & M. Petzell (2013) The impact of sociolinguistic factors on the usage of the pre-prefix in some G languages. Paper presented at *Fifth International Conference on the Bantu Languages, June 12-15, Paris*.
- Baka, J. (2000) *L'adjectif en bantu*. Bruxelles: Université libre de Bruxelles, thèse de doctorat.
- Baka, J. (2005) Pourquoi une forme de type \*-pái “nouveau” en protobantou. In Bostoen, K. & J. Maniacky (eds.), *Studies in African Comparative Linguistics, with Special Focus on Bantu and Mande*, 139-146. Tervuren: Royal Museum for Central Africa.
- Bastin, Y. (1989) Les déverbatifs bantous en -e. *Journal of African Languages and Linguistics* 11: 151-174.
- Bastin, Y. & A. Coupez & E. Mumba & T. C. Schadeberg (eds.) (2002) *Bantu Lexical Reconstructions 3*. Tervuren: Royal Museum for Central Africa, online database: [linguistics.africamuseum.be/BLR3.html](http://linguistics.africamuseum.be/BLR3.html) (last accessed November 28, 2017).
- Blanchon, J. A. (1997) Les formes nominales de citation a préfixe haut en pounou (bantu B43). *Journal of African Languages and Linguistics* 18(2): 129-138.
- Blanchon, J. A. (1998) Semantic/pragmatic conditions on the tonology of the Kongo noun-phrase: a diachronic hypothesis. In Hyman, L. M. & C. W. Kisseberth (eds.), *Theoretical aspects of Bantu tone*, 1-32. Stanford: Center for the Study of Language and Information, Stanford University.
- Bokamba, G. D. (1971) Specificity and definiteness in Dzamba. *Studies in African Linguistics* 2: 217-237.
- Buell, L. (2009) The distribution of the Nguni augment: a review. Paper presented at *Bantu Augment Workshop, Leiden University, June 17, Leiden*.
- Carter, H. (1963) Coding, style and the initial vowel in north Rhodesian Tonga: a psycholinguistic study. *African Language Studies* 4: 1-42.
- Cinque, G. (2005) Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36(3): 315-332.
- Corbett, G. G. (1979) The agreement hierarchy. *Journal of Linguistics* 15: 203-224.
- Corbett, G. G. (1991) *Gender*. Cambridge: Cambridge University Press.
- Crane, T. M. & L. M. Hyman & S. Nsielanga Tukumu (2011) *A Grammar of Nzadi (B.865) A Bantu Language of the Democratic Republic of Congo with an Appendix of Proto-Bantu - Nzadi Sound Correspondences by Clara Cohen*. Berkeley: University of California Publications.
- Creissels, D. (2006) *Syntaxe générale, une introduction typologique. 2. La phrase*. Paris: Lavoisier.
- Creissels, D. (2009) Construct forms of nouns in African languages. In Austin, P. K. & O. Bond & M. Charette & D. Nathan & P. Sells (eds.), *Proceedings of Conference on Language Documentation & Linguistic Theory 2*, 73-82. London: SOAS.
- Creissels, D. (2014) The 'new adjectives' of Tswana. In Simone, R. & F. Masini (eds.), *Word Classes: Nature, Typology and Representations*, 75-94. Amsterdam; Philadelphia: John Benjamins.

- Croft, W. (2001) *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
- Daniel, M. & E. A. Moravcsik (2005) Associative plurals. In Dryer, M. S. & M. Haspelmath & D. Gil & B. Comrie (eds.), *The World Atlas of Language Structures*, 150-153. Oxford: Oxford University Press.
- De Blois, K. F. (1970) The Augment in the Bantu languages. *Africana Linguistica* 4: 85-165.
- Devos, M. (2008) *A Grammar of Makwe*. München: Lincom Europa.
- Dimmendaal, G. (2011) *Historical Linguistics and the Comparative Study of African Languages*. Amsterdam; Philadelphia: John Benjamins.
- Doke, C. M. (1927) The significance of class 1a of Bantu nouns. In Boas, F. & O. H. Dempwolff & G. Panconcelli-Calzia & A. Werner & D. Westermann (eds.), *Festschrift Meinhof: Sprachwissenschaftliche und andere Studien*, 196-203. Hamburg: : Kommissionsverlag von L. Friederichsen.
- Doke, C. M. (1997) *Textbook of Zulu Grammar. Sixth Edition*. Cape Town: Maskew Miller Longman.
- Doneux, J. L. (1967) Données sur la classe 15 nominale en bantou. *Africana Linguistica* 3: 1--22.
- Ernst, U. (1992) *Esquisse grammaticale du kako*. Yaoundé: Summer Institute of Linguistics.
- Forges, G. (1983) La classe de l'infinitif en bantou. *Africana Linguistica* 9: 257-264.
- Givón, T. (1974) Syntactic change in Lake Bantu: a rejoinder. *Studies in African Linguistics* 5(1): 117-139.
- Givón, T. (2001) *Syntax : An Introduction*. Amsterdam; Philadelphia: John Benjamins.
- Good, J. (2012) How to become a "Kwa" noun. *Morphology* 22(2): 293-335.
- Greenberg, J. H. (1963) Some universals of grammar with particular reference to the order of meaningful elements. In Greenberg, J. H. (ed.), *Universals of Language*. Cambridge: M.I.T. Press.
- Greenberg, J. H. (1978) How does a language acquire gender markers. In Greenberg, J. H. & C. A. Ferguson & A. Moravcsik (eds.), *Universals of Human Language, Volume 3 Word Structure*, 47-82. Stanford: Stanford University Press.
- Gregersen, E. (1967) *Prefix and pronoun in Bantu*. [Baltimore,: Published at the Waverly Press by Indiana University, Bloomington.
- Grégoire, C. (1975) *Les locatifs en bantou*. Tervuren: Musée royal de l'Afrique centrale.
- Grégoire, C. & B. Janssens (1999) L'augment en bantou du nord-ouest. In Hombert, J.-M. & L. Hyman (eds.), *Bantu Historical Linguistics: Theoretical and Empirical Perspectives*, 413-429. Stanford: CSLI.
- Hadermann, P. (1999) Les formes nomino-verbales de classes 5 et 15 dans les langues bantoues du Nord-Ouest. In Hombert, J.-M. & L. M. Hyman (eds.), *Bantu historical linguistics: theoretical and empirical perspectives*, 431-471. Stanford: CSLI.
- Harjula, L. (2004) *The Ha Language of Tanzania: Grammar, Texts and Vocabulary*. Cologne: Rüdiger Köppe.
- Hawkins, J. A. (1983) *Word Order Universals*. New York: Academic Press.
- Heath, T. (1998) *A Preliminary Grammar Sketch of the Meka Noun and Verb Morphology*. Yaoundé: SIL Cameroon.
- Hendrikse, A. P. (1990) Number as a categorizing parameter in Southern Bantu: an exploration in cognitive grammar. *South African Journal of African Languages* 10(4): 384-400.
- Hulstaert, G. (1965) *Grammaire du lomongo. Deuxieme partie, Morphologie*. Tervuren: Musée royal de l'Afrique centrale.
- Hulstaert, G. (1966) *Grammaire du lomongo. Troisième partie, Syntaxe*. Tervuren: Musée royal de l'Afrique centrale.

- Hyman, L. M. (1972) A Phonological Study of Fe'fe'-Bamileke. *Studies in African Linguistics*. Supplement 4.
- Hyman, L. M. (2003) Basaá (A43). In Nurse, D. & G. Philippson (eds.), *The Bantu Languages*, 257-282. London New York: Routledge.
- Hyman, L. M. & F. X. Katamba (1993) The Augment in Luganda: Syntax or Pragmatics? In Mchombo, S. A. (ed.), *Theoretical Aspects of Bantu Grammar*, 209-256. Stanford: CSLI.
- Kamanda Kola, R. (1993) A propos de la tonalité en zamba. *Afrikanistische Arbeitspapiere* 33: 83-103.
- Kamanda Kola, R. (1994) Notes sur l'augment en zamba. *Annales Aequatoria* 15: 399-409.
- Kamba-Muzenga, J.-G. (2003) *Substitutifs et possessifs en bantou*. Paris: Peeters.
- Kasonde, A. R. M. (2009) *Phonologie et morphologie de la langue bemba*. München: LINCOM.
- Kavari, J. U. & L. Marten & J. v. D. Wal (2012) Tone cases in Otjiherero: head-complement relations, linear order, and information structure. *Africana Linguistica* 18: 315-353.
- Kenstowicz, M. (2008) On the Origin of Tonal Classes in Kinande Noun Stems. *Studies in African Linguistics* 37(2): 115-151.
- Leitch, M. (2003) Babole (C101). In Nurse, D. & G. Philippson (eds.), *The Bantu Languages*, 392-421. London New York: Routledge.
- Maganga, C. & T. C. Schadeberg (1992) *Kinyamwezi: Grammar, Texts, Vocabulary*. Cologne: Rüdiger Köppe.
- Maho, J. F. (1999) *A comparative study of Bantu noun classes*. Göteborg: Acta Universitatis Gothoburgensis.
- Meeussen, A. E. (1959) *Essai de grammaire rundi*. Tervuren: Musée royal du Congo belge.
- Meeussen, A. E. (1967) Bantu Grammatical Reconstructions. *Africana Linguistica* 3: 79-121.
- Meeuwis, M. (2010) *A Grammatical Overview of Lingála* München: Lincom.
- Möhlig, W. J. G. & L. Marten & J. U. Kavari (2002) *A Grammatical Sketch of Herero*. Cologne: Rüdiger Köppe.
- Mous, M. (2003) Nen (A44). In Nurse, D. & G. Philippson (eds.), *The Bantu Languages*, 283-306. London; New York: Routledge.
- Mudzingwa, C. & M. Kadenge (2013) An analysis of the ghost augment in chiShona. *South African Journal of African Languages* 33(1): 87-93.
- Nicolle, S. (2012) Semantic-Pragmatic Change in Bantu -no Demonstrative Forms. *Africana Linguistica* 18: 193-234.
- Nicolle, S. (2013) *A grammar of Digo: A Bantu language of Kenya and Tanzania*. Dallas: SIL International.
- Nicolle, S. (2014) Discourse Functions of Demonstratives in Eastern Bantu Narrative Texts. *Studies in African Linguistics* 43(2): 125-144.
- Nsuka-Nkutsi, F. (1982) *Les structures fondamentales du relatif dans les langues bantoues*. Tervuren: Musée Royal de l'Afrique Centrale.
- O'Connor, K. M. & C. Patin (2015) The syntax and prosody of apposition in Shingazidja. *Phonology* 32(1): 111-145.
- Ondo Mebiame, P. (2001) L'augment en fang ntumu. *Revue gabonaise des sciences du langage* 2: 60-77.
- Patin, C. (2017) The Augment in Shingazidja. Paper presented at *the SOAS-UGent Workshop on Parametric Approaches to Morphosyntactic Variation in Bantu*, 31 March 2017, Ghent University.
- Petzell, M. (2008) *The Kagulu Language of Tanzania: Grammar, Texts and Vocabulary*. Cologne: Rüdiger Köppe.
- Polak-Bynon, L. (1965) L'expression des ordinaux dans les langues bantoues. *Africana Linguistica* 2: 127-159.

- Rijkhoff, J. (2002) *The Noun Phrase*. Oxford: Oxford University Press.
- Rijkhoff, J. (2008) Descriptive and discourse-referential modifiers in a layered model of the noun phrase. *Linguistics* 46(4): 789-829.
- Schadeberg, T. C. (1986) Tone cases in Umbundu. *Africana Linguistica* 10: 423-447.
- Schadeberg, T. C. (2003) Derivation. In Nurse, D. & G. Philippson (eds.), *The Bantu Languages*, 71-89. London; New York: Routledge.
- Segerer, G. (2015) Les classes nominales en keerak. In Creissels, D. & K. Pozdniakov (eds.), *Les classes nominales dans les langues atlantiques*, 103-146. Cologne: Rüdiger Köppe.
- Segerer, G. & S. Flavier (2011-2016) *RefLex: Reference Lexicon of Africa, Version 1.1*. Paris, Lyon: CNRS, <http://reflex.cnrs.fr/>.
- Stappers, L. (1964) *Morfologie van het Songye*. Tervuren: Koninklijk Museum voor Midden-Afrika.
- Stappers, L. (1965) Het hoofdtelwoord in de Bantoe-talen. *Africana Linguistica* 2: 177-199.
- Ström, E.-M. (2013) *The Ndengeleko Language of Tanzania*. Gothenburg: University of Gothenburg, PhD dissertation.
- Taylor, C. (1985) *Nkore-Kiga*. London: Croom Helm.
- Torrend, J. (1891) *A Comparative grammar of the South-African Bantu languages, comprising those of Zanzibar, Mozambique, the Zambezi, Kafirland, Benguela, Angola, the Congo, the Ogowe, the Cameroons, the Lake Regions*. London: Kegan Paul, Trench, Trübner & Co.
- Twilingiyimana, C. (1984) *Éléments de description du doko*. Tervuren: Musée royal de l'Afrique centrale.
- Valinande, N. K. (1984) *The structure of Kinande*. Washington DC: Georgetown University, PhD dissertation.
- Van de Velde, M. (2005) The order of noun and demonstrative in Bantu. In Bostoen, K. & J. Maniacky (eds.), *Studies in African Comparative Linguistics, with Special Focus on Bantu and Mande*, 425-441. Tervuren: Royal Museum for Central Africa.
- Van de Velde, M. (2006a) The alleged class 2a prefix bò in Eton, a plural word. Paper presented at *31st Annual Meeting of the Berkeley Linguistics Society, February 2005*, Berkeley, <https://halshs.archives-ouvertes.fr/halshs-00603470>.
- Van de Velde, M. (2006b) Multifunctional Agreement Patterns in Bantu and the Possibility of Genderless Nouns. *Linguistic Typology* 10: 183-221.
- Van de Velde, M. (2008) *A grammar of Eton*. Berlin: Mouton De Gruyter
- Van de Velde, M. (2009) Agreement as a grammatical criterion for proper name status in Kirundi. *Onoma* 44: 219-241.
- Van de Velde, M. (2013) The Bantu connective construction. In Carlier, A. & J.-C. Verstraete (eds.), *The Genitive*, 217-252. Amsterdam: John Benjamins.
- Van de Velde, M. (2017) The augment as a construct form marker in Eton relative clause constructions. In Atindogbe, G. & R. Grollemund (eds.), *Relative Clauses in Cameroonian Languages: Structure, Function and Semantism*, tba. Berlin: De Gruyter Mouton.
- Van de Velde, M. & O. Ambouroué (2011) The grammar of Orungu proper names. *Journal of African Languages and Linguistics* 32: 113-141.
- van den Eynde, K. & P. Mufuta (1994) Le système de référence primaire du ciluba (luba-Kasai): analyse en traits. *International Journal of Lexicography* 7: 90-105.
- Van der Wal, J. (2006) Predicative Tone Lowering in Makhuwa. *Linguistics in the Netherlands* 23: 224-236.
- Van Langendonck, W. (2007) *Theory and Typology of Proper Names*. Berlin: Mouton de Gruyter.
- Van Leynseele, H. (1977) *An outline of Libinza grammar*. Leiden: Rijksuniversiteit Leiden, MA thesis.

- von Staden, P. M. S. (1973) The initial vowel of the noun in Zulu. *African Studies* 32(3): 163-181.
- Wald, B. (1975) Animate concord in Northeast Coastal Bantu: Its linguistic and social implications as a case of grammatical convergence. *Studies in African Linguistics* 6(3): 267-314.
- Wega Simeu, A. (2012) *Grammaire descriptive du pólri. Eléments de phonologie, de morphologie et de syntaxe*. Yaounde: University of Yaounde I, thèse de doctorat.
- Welmers, W. E. (1973) *African language structures*. Berkeley, Los Angeles & London: University of California Press.
- Wilkins, W. & A. Kimenyi (1975) Strategies in constructing a definite description: some evidence from Kinyarwanda. *Studies in African Linguistics* 6(2): 151-170.
- Ziervogel, D. & E. J. Mabuza (1976) *A Grammar of the Swati Language: Siswati*. Pretoria: J.L. Van Schaik.

---

<sup>i</sup> These observations are based on data from a sample of fifteen languages extracted from the RefLex database (Segerer & Flavier 2011-2016). The sample is restricted to sources of at least a thousand lexical entries, created by specialists of the language (i.e. excluding lexical surveys) and counts exclude the even numbered classes and everything up from class 13 (included). Percentages of nouns belonging to class 1 in the (North)west include: 2% (Eton A71), 1% (Ngom/Koya B22), 5% (Gevia B301) and 2% (Yombe H16c). More to the East we find: 14% (Lega D25), 13% (Nyamwezi F22), 16% (Ndamba G52), 15% (Makonde P23) and 11% (Tswana S31). The north-western A80 languages are a notable exception to this trend with an average of 20% of nouns belonging to class 1 in three languages (Kol A832; Njyem A84 and Mekaa A83). In two of these languages, class 9 is marginal, which may or may not be a coincidence. For class 7, differences in size between western and eastern languages are less dramatic. Typical figures are 34% (Njyem A84) versus 16% (Makonde P23).

<sup>ii</sup> These subclasses are not to be confused with the notion of ‘subgenders’ in the typological literature, which involve minimally distinct agreement patterns (Corbett 1991: 163).

<sup>iii</sup> Restrictive relative clauses always require the augment in these languages, whatever the type of modification they provide. This may be due to the historical or current presence of a demonstrative relativiser.

<sup>iv</sup> Both authors, Leitch (2003) and Kamanda Kola (1994), describe the distinction in terms of definiteness, but again this analysis seems to be more inspired by the French translation equivalents of these examples than by the Babole and Zamba data. Note in this respect that, if the augment were a definiteness marker in (44b), we would have expected it to agree with and be prefixed to the dependent noun ‘woman’. Instead, it is prefixed to the connective relator, and it agrees with the head noun ‘navel’.

<sup>v</sup> The linkers that are obligatorily used with a subset of adnominal modifiers in Southern Bantu languages, such as Tswana S31 (Creissels 2006: 75, 2014) are another instance of the same general phenomenon. They originate in demonstratives and can be used to turn sequences of proclitic subject marker plus descriptive predicate into adnominal modifiers, arguably via a stage of nominalisation.