

Noun classification, definiteness, number, possession

1. Noun classification

Classifier systems of the type encountered in languages of East Asia or of the Pacific are extremely rare in Africa. A system of genitival classifiers has been described in the Ubangian language Dongo-ko, and a system of numeral classifiers has been described in the Cross-River language Kana, but these are quite exceptional cases.

By contrast, one commonly finds in Africa noun classification systems in which nouns are divided into several subsets on the basis of their agreement properties with noun dependents, verbs and pronouns. Two types of noun classification systems based on agreement are common among African languages:

- Systems with two genders in which the sex distinction (masculine vs. feminine) is relevant to gender assignment are found in all branches of Afroasiatic¹, in some branches of Nilo-Saharan, in Kadu (a group of languages spoken in the Nuba mountains which may belong to Nilo-Saharan or constitute a genetic isolate), and among Khoisan languages, in Khoe languages, Sandawe, Kwadi, and Hadza. A third gender similar to the Indo-European neuter has been reported to exist in Eastern Nilotic languages and in Khoe languages.
- Another type of gender system, with a much higher number of genders, and in which the sex distinction plays no role, is encountered in most major branches of Niger-Congo.² Its most typical representatives are found among Atlantic languages and Bantu languages, i.e. in two areas very distant from each other. This type of gender system is discussed in Appendix 1.

Noun classification systems typologically close to the Niger-Congo type are found in North and South Khoisan. They however differ from Niger-Congo systems in that the number of classes never exceeds 5, and class membership generally does not manifest itself in noun morphology, but only in agreement.

A few Niger-Congo languages (e.g. Ijo, the Ubangian language Zande, the Mande language Jo) have a *masculine vs. feminine* distinction in pronouns, which however is not involved in the agreement mechanisms typical of gender systems. The languages in question do not correspond to any grouping definable in genetic or geographic terms.

¹ However, among Chadic languages, the gender distinction is not general.

² The absence of any vestige of the Niger-Congo noun classification system in Mande languages is one of the main reasons to question the inclusion of Mande in the Niger-Congo phylum.

2.Referentiality and definiteness

Languages with and without definite articles are encountered virtually in all language families and in all parts of the African continent, and there is often clear evidence that they are cognate with demonstratives. This confirms that the grammaticalization process ‘demonstrative → definite article’ is very frequent in the evolution of languages,³ but also that definite articles are relatively unstable, and tend to be affected by processes leading to their loss or to a change in their status. There is a very general tendency of definite articles proper towards expanding their use to include both definite determination and non-definite referential uses, giving rise to what Greenberg called ‘stage II of the definite article’.

Articles at an advanced stage of their evolution are particularly common in Africa, and the relatively high proportion of African languages with drastically eroded “stage-II articles” is remarkable: in a number of African languages, “stage-II articles” manifest themselves only through a change in tone at the beginning or at the end of the word they are attached to, which results from the erosion of former prefixes or suffixes.

For example, Kita Maninka (Mande) has a ‘definite’ form of nouns marked by tone only. In some contexts, in particular, in negative clauses – ex. (1a-b), its use involves semantic distinctions typically expressed by articles, but in other contexts the ‘indefinite’ form is impossible, and the definite form has the status of a default form of nouns that by itself does not imply any semantic specification – ex. (1c-d).

- (1) a. *Ń mán wóri dí Músá mà.*
 1SG PFV.POS money:DEF give Moussa to⁴
 ‘I did not give the money to Moussa.’
- b. *Ń mán wóri dí Músá mà.*
 1SG PFV.POS money give Moussa to
 ‘I did not give money to Moussa.’
- c. *Ń dí wóri dí Músá mà.*
 1SG PFV.POS money:DEF give Moussa to
 ‘I gave (the) money to Moussa.’
- d. **Ń dí wóri dí Músá mà.*
 1SG PFV.POS money give Moussa to

In a synchronic description of Kita Maninka, the definite form of nouns can be described as underlyingly including a floating low tone in addition to its lexical tone pattern, and comparative data shows that, historically, this floating low tone results

³ Some languages provide however evidence that definite articles may originate from possessives too.

⁴ The abbreviations used in the glosses are as follows: CL = noun class, DEF = definite, DEM = demonstrative, DIST = distal, F = feminine, GEN = genitive, H = head-marking, LNK = linker, M = masculine, PFV = perfective, POS = positive, POSTP = postposition, PRF = perfect, PRO = pronoun, PRS = present, REL = relativizer, SG = singular.

from the erosion of a cliticized form of a demonstrative still found in Kita Maninka as ò ‘that’.

1.3. Number

In the languages of the world, plural marking restricted to a narrow range of nouns is not uncommon, but most African languages have bound morphemes encoding plurality whose use is not restricted to nouns occupying a relatively high position in the animacy hierarchy.

The total lack of plural markers is however illustrated by the Western Benue-Congo language Igbo. In this language, the two nouns meaning ‘child’ and ‘person’ have suppletive plural forms, but with nouns that are not compounds having ‘person’ or ‘child’ as their first formant, plurality can be expressed only by adding numerals or quantifiers such as ‘several’, ‘a few’, ‘many’, etc. Such a situation is found also in some languages of the Chadic family (Gwandara, Pero), but on the whole, it is rather exceptional in Africa.

The same can be said of systems of plural markers restricted to a narrow range of nouns (mainly human and animate); such a situation is only sporadically found among African languages (for example, in some languages of the Chadic family).

A common number-marking system attested in a wide range of Nilo-Saharan languages involves a three-way distinction between nouns with an overt plural marker and a morphologically unmarked singular form, those with an overt singular marker and a morphologically unmarked plural form, and those morphologically marked both in the singular and the plural, as in the Maban language Aiki (Runga):

(2)	<i>Singular</i>	<i>Plural</i>	
	àyó-k	àyó	‘leaf’
	k̀l̀l̀	k̀l̀l̀-t	‘snake’
	d̀d̀d̀-í	d̀d̀d̀-ú	‘leg’

In such systems, collective entities such as ‘leaf’, ‘hair’ or ‘tooth’ or words referring to items naturally occurring in pairs, such as ‘shoe’, ‘eye’ or ‘wing’, tend to be morphologically unmarked in the plural; the corresponding singular expresses an individuated item from a collective or from a pair.

As regards the use of plural markers, two opposite tendencies emerge among African languages, which are not bound to any particular genetic or geographical grouping, but rather seem to correlate both with the morphological nature of plural markers and with the presence vs. absence of a gender system:

- Languages devoid of a gender system frequently have a single plural marker with the morphological status of a phrasal affix, and such plural markers tend to be used on a ‘pragmatic’ basis, i.e. to be employed only when plurality is both communicatively relevant and not implied by the context, at least in the case of nouns that do not refer to persons.
- Languages that have gender generally have a morphologically complex plural marking, characterized by a fusion of gender and number markers, and variations in gender and number manifest themselves through morphemes affixed to the head noun and to (some of) its modifiers, in an agreement relationship. In these languages, plural marking tends to function on a ‘semantic’

basis, which means that plural markers tend to be present in every noun phrase referring to a plurality of individuals, irrespective of their communicative relevance.

Extreme cases of morphologically complex number marking are encountered in an area including the Eastern-Sudanic branch of Nilo-Saharan and all branches of the Afroasiatic phylum.

Among the possible types of number systems, only the most common ones are well represented among African languages. 'Greater plural' is however relatively common among Atlantic and Bantu languages, i.e. among the Niger-Congo languages that have the Niger-Congo noun classification system in its most typical form; in these languages, singular and plural are distinguished by class affixes, and some noun stems at least can combine with two distinct class affixes to express a distinction between ordinary plural and greater plural, as in Southern Sotho *pere* (cl. 9) 'horse, *li-pere* (cl. 10) 'horses, *ma-pere* (cl. 6) 'great many horses'.

1.4. The genitival modifier

The genitival construction may involve head-marking, with a special 'construct form' of the head noun or possessive affixes attached to the head noun, or dependent-marking, with a genitive marker attached to the genitival modifier.

Case marked genitives are relatively common in Africa, even in languages devoid of case contrast between subject and object. Genitival constructions involving obligatory possessive affixes even in the presence of a noun phrase in genitive function are not very common, but they are found in different language families and in different parts of the African continent, and the same can be said of genitival constructions involving a construct form of the head noun. On 'construct forms', see Appendix 2.

Cross-linguistically, gender-number agreement of genitive NPs with their head is not very common, but it occurs in the most typical Niger-Congo noun class systems (particularly in Bantu languages).

A number of African languages have more than one possible way of combining a noun with a genitival modifier, most commonly with a distinction in meaning so that the variant with more morphological material (genitive markers or possessive affixes) is used with 'non-intimate' (or 'alienable') types of relations, and that with less morphological material with 'intimate' (or 'inalienable') types of relations. This type of distinction is particularly common in some language groups (e.g. Mande languages), but it is not really bound to particular families or areas.

For example, in Mandinka (Mande), the genitive precedes its head. It may be simply juxtaposed ('inalienable' construction), or marked by the postposition *la* ('alienable' construction). The main regularities are that:

- (a) With inanimate genitives, the morphologically unmarked construction is usual, and the construction with the postposition *la* is found only if the genitival relationship is the transposition of a subject-verb relationship.
- (b) With animate genitives, the construction with the postposition *la* constitutes the default choice, and the morphologically unmarked construction is found mainly with head nouns referring to body parts or kinship relationships, or

when the genitival relationship is the transposition of an object-verb relationship.

- (3) a. *búŋo kódo* 'the money of the house'
house:DEF money:DEF
- b. *kewó la kódo* 'the man's money'
man:DEF POSTP money:DEF
- c. *kewó kuŋô* 'the man's head'
man:DEF head:DEF
- d. *díndíŋo mamáúso* 'the child's grandmother'
child:DEF grandmother:DEF

Appendix 1. Niger-Congo noun class systems: prototype and variations

A1.1. Introduction

In general typological terms, Niger-Congo noun class systems are a particular variety of gender systems, since the notion of noun class in the description of Niger-Congo languages refers to a division of nouns into subsets manifested in their behavior in agreement mechanisms.

This section discusses the definition of a prototype of Niger-Congo noun class systems, with the object of helping to put them in the broader perspective of a general typology of noun classification systems and of the grammaticalization processes in which they are involved. Illustrations are taken from Tswana, a southern Bantu language whose noun classification system stands very close to the prototype.

A1.2. Niger-Congo noun class systems as very grammaticalized systems of agreement in which gender marking cannot be dissociated from number marking

The essential feature of Niger-Congo classification systems is that

- (a) noun forms divide into subsets (noun classes) according to their behavior in agreement mechanisms observed in the formation of noun phrases by combining a head noun with various types of modifiers, in the use of pronouns, and in the indexation of arguments on the verb,
- (b) the forms involved in these agreement mechanisms (nouns, noun modifiers, pronouns and verbs) include affixes (class markers) that determine their behavior as controllers or targets of agreement, and
- (c) the classification is a lexicalized classification of nouns, and not a classification of referents directly, in the sense that it is not possible to change the class of a

noun in order to emphasize particular semantic features of the same referent, as typically observed in classifier systems.⁵

Semantically, the most obvious function of noun class markers (but not the only one – see below) is the expression of number (singular vs. plural). An important feature of Niger-Congo classification systems is the absolute impossibility to isolate morphemes expressing number independently from gender. Moreover, there is generally no one-to-one correspondence between agreement classes and class markers in the singular and in the plural. This is the reason why many descriptions of Niger-Congo class systems do not emphasize the possibility of dividing noun *lexemes* into *genders*, but rather start from a division of noun *forms* into *classes* in which the singular form and the plural form of a given noun are treated as two distinct units; in this approach, a gender may be subsequently defined as a couple of classes that include the singular and plural forms of the same lexemes. For example, in Tswana, *mosadi* pl. *basadi* ‘woman’ belongs to gender 1/2, which means that its singular form *mosadi* belongs to the agreement class labeled ‘1’, and its plural form *basadi*, to the agreement class labeled ‘2’.⁶

In the description of Niger-Congo noun class systems, the regularity of agreement generally makes it easy to establish the number of classes into which noun forms divide (i.e., the number of possible agreement patterns); by contrast, the idiosyncrasies shown by many nouns in the singular-plural correspondence and the variations observed in plural formation often make it very difficult to decide how many genders must be recognized, if genders are defined as sets of nominal lexemes with the same agreement properties both in the singular and in the plural.

(1) *the division of Tswana noun forms into 12 classes on the basis of the agreement between nouns and adjectives in the attributive construction*⁷

a. cl. 1	<i>mosadi yo moša</i>	mù-sàdì jó mù-ǰá	‘new woman’
b. cl. 2	<i>basadi ba baša</i>	bà-sàdì bá bà-ǰá	‘new women’
c. cl. 3	<i>molemo o moša</i>	mù-lìmò ó mù-ǰá	‘new medicine’
d. cl. 4	<i>melemo e meša</i>	mì-lìmò é mì-ǰá	‘new medicines’
e. cl. 5	<i>lesaka le leša</i>	lì-sàkà lé lì-ǰá	‘new cattle kraal’
f. cl. 6	<i>maraka a maša</i>	mà-ràkà á mà-ǰá	‘new cattle kraals’
g. cl. 7	<i>sekolo se seša</i>	sì-kólò sé sì-ǰá	‘new school’
h. cl. 8-10	<i>dikolo tse dintšha</i> <i>dikgosi tse dintšha</i>	di-kólò tsé díh-tʰá di-qʰòsì tsé díh-tʰá	‘new schools’ ‘new chiefs’
i. cl. 9	<i>kgosi e ntšha</i>	qʰòsì é ò-tʰá	‘new chief’
j. cl. 11	<i>lokwalò lo loša</i>	lù-kʰálò ló lò-ǰá	‘new book’

⁵ In Niger-Congo noun class systems, a given nominal lexeme may be compatible with several class markers, but the variations express changes in the referential meaning (see below), which is an entirely different semantic operation.

⁶ The numbering of classes used in Bantu studies is based on the correspondence with the reconstructed classes of Proto-Bantu; for example, in the description of Tswana, ‘cl. 8-10’ refers to a class that historically results from the merger of Proto-Bantu classes 8 and 10, and the absence of classes 12 and 13 means that no Tswana class is the reflex of Proto-Bantu classes 12 and 13.

⁷ This construction involves an obligatory linker expressing class agreement too. Historically, this linker is the reflex of a former demonstrative that has lost its semantic content and has become a purely formal element of the construction.

k. cl. 14	<i>bojang jo boša</i>	bùdʒàŋ dʒó bù-ǰá	‘new grass’
l. cl. 15-17	<i>go lema mo goša</i>	χù-limà mó χù-ǰá	‘new way of cultivating’

(2) *Tswana noun forms of cl. 1, 5, 7 and 9 (or the corresponding plural, if meaning requires it) combined with various types of modifiers (genitives, demonstratives, etc.)*

a. <i>mosadi wa motse</i>	mù-sádì w-á-mú-tsi	‘woman of the village’
<i>lekau la motse</i>	lì-káù l-á-mú-tsi	‘boy of the village’
<i>sefofu sa motse</i>	sì-fòfù s-á-mú-tsi	‘blind person of the village’
<i>ngaka ya motse</i>	ŋàkà j-á-mú-tsi	‘doctor of the village’
b. <i>mosadi yole</i>	mù-sádì jó-lé	‘that woman’
<i>lekau lele</i>	lì-káù lé-lé	‘that boy’
<i>sefofu sele</i>	sì-fòfù sé-lé	‘that blind person’
<i>ngaka ele</i>	ŋàkà é-lé	‘that doctor’
c. <i>mosadi ofe?</i>	mù-sádì ú-ǰí	‘which woman?’
<i>lekau lefe?</i>	lì-káù lí-ǰí	‘which boy?’
<i>sefofu sefe?</i>	sì-fòfù sí-ǰí	‘which blind person?’
<i>ngaka efe?</i>	ŋàkà í-ǰí	‘which doctor?’
d. <i>mosadi osele</i>	mù-sádì ú-sìlí	‘another woman’
<i>lekau lesele</i>	lì-káù lí-sìlí	‘another boy’
<i>sefofu sesele</i>	sì-fòfù sí-sìlí	‘another blind person’
<i>ngaka esele</i>	ŋàkà í-sìlí	‘another doctor’
e. <i>mosadi ope</i>	mù-sádì ú-pé	‘no woman’
<i>lekau lepe</i>	lì-káù lí-pé	‘no boy’
<i>sefofu sepe</i>	sì-fòfù sí-pé	‘no blind person’
<i>ngaka epe</i>	ŋàkà í-pé	‘no doctor’
f. <i>basadi botlhe</i>	bà-sádí b-ótl ^h é	‘every woman’
<i>makau otlhe</i>	mà-káú ótl ^h é	‘every boy’
<i>difofu tsotlhe</i>	dì-fòfù ts-ótl ^h é	‘every blind person’
<i>dingaka tsotlhe</i>	dì-ŋàkà ts-ótl ^h é	‘every doctor’
g. <i>mosadi mongwe</i>	mù-sádí mú-ŋ ^w ì	‘one woman’
<i>lekau lengwe</i>	lì-káú lí-ŋ ^w ì	‘one boy’
<i>sefofu sengwe</i>	sì-fòfù sì-ŋ ^w ì	‘one blind person’
<i>ngaka nngwe</i>	ŋàkà ñ-ŋ ^w ì	‘one doctor’
h. <i>basadi ba le babedi</i>	bà-sádí bá-lí bá-bèdí	‘two women’
<i>makau a le mabedi</i>	mà-káú á-lí má-bèdí	‘two boys’
<i>difofu di le pedi</i>	dì-fòfù dí-lí pèdí	‘two blind persons’
<i>dingaka di le pedi</i>	dì-ŋàkà dí-lí pèdí	‘two doctors’

i. *mosadi yo ke mo thusitseng maabane*

mù-sádì jò kí-mù-t^húsítsèǰ máàbání
 CL1-blind CL7.LNK 1SG-CL7-help:PRF:REL yesterday
 ‘the woman I helped yesterday’

lekau le ke le thusitseng maabane

lì-káù lé kí-lì-t^húsítsèǰ máàbání
 CL5-boy CL5.LNK 1SG-CL7-help:PRF:REL yesterday
 ‘the boy I helped yesterday’

sefofu se ke se thusitseng maabane

sì-fòfù sé kí-sì-t^húsítsèǰ máàbání
 CL7-blind CL7.LNK 1SG-CL7-help:PRF:REL yesterday
 ‘the blind person I helped yesterday’

ngaka e ke e thusitseng maabane

ǰàkà é kí-ì-t^húsítsèǰ máàbání
 [CL9]doctor CL9.LNK 1SG-CL9-help:PRF:REL yesterday
 ‘the doctor I helped yesterday’

(3) *Obligatory presence of class markers on every element of complex Tswana noun phrases*

a. *mosadi yo moleele yo montsho yo o opelang yole*

mù-sádì jò mù-léèlé jò mú-̀nts^hù
 CL1-woman CL1.LNK CL1-tall CL1.LNK CL1-black
 jò ú-̀pélàǰ jò-lé
 CL1.LNK CL1-sing:PRS:RELDEM.CL1-DIST
 ‘this tall woman with dark complexion who is singing’

b. *lekau le leleele le lentsho le le opelang lele*

lì-káù lé lì-léèlé lé lí-̀nts^hù
 CL5-boy CL5.LNK CL5-tall CL5.LNK CL5-black
 lé lí-̀pélàǰ lé-lé
 CL5.LNK CL5-sing:PRS:RELDEM.CL5-DIST
 ‘this tall boy with dark complexion who is singing’

(4) *Class markers affixed to verbs in subject and object marker function, and free pronouns marked for class*

a. *Mosadi o lapile* mù-sádì ú-lápile ‘The woman is tired’
 CL1-woman CL1-be_tired:PRF

Lekau le lapile lì-káù lí-lápile ‘The boy is tired’
Sefofu se lapile sì-fòfù sí-lápile ‘The blind person is tired’
Ngaka e lapile ǰàkà í-lápile ‘The doctor is tired’

b. *Ke mo thusitse* kì-mù-t^húsítsè ‘I helped her (the woman)’
 1SG-CL1-help:PRF

<i>Ke le thusitse</i>	kì-lí-t ^h úsítsè		‘I helped him (the boy)’
<i>Ke se thusitse</i>	kì-sí-t ^h úsítsè		‘I helped him/her (the blind person)’
<i>Ke e thusitse</i>	kì-í-t ^h úsítsè		‘I helped him (the doctor)’
c. <i>Ke bua le ene</i>	kì-búà	lí-ènéné	‘I am speaking with her (the woman)’
	1SG-speak:PRS	with-CL1.PRO	
<i>Ke bua le lone</i>	kì-búà	lí-l-òné	‘I am speaking with him (the boy)’
<i>Ke bua le sone</i>	kì-búà	lí-s-òné	‘I am speaking with him/her (the blind person)’
<i>Ke bua le yone</i>	kì-búà	lí-j-òné	‘I am speaking with him (the doctor)’

A1.3. Complexities of Niger-Congo noun class systems at morphological level

After examining characteristics of Niger-Congo languages that are crucial in the definition of a prototype of Niger-Congo noun class systems, we now examine some other features particularly relevant from the point of view of typological comparison, and with respect to which some variation is observed.

As regards form, all Niger-Congo noun class systems depart more or less from the ideally simple situation in which all noun forms belonging to a given agreement class invariably show an overt and non ambiguous class marker identical for all nouns that have the same agreement properties, and in which a class marker identical to that attached to nouns is simply repeated on each form expressing class agreement with nouns. The Tswana examples above make it clear that some degree of irregularity can always be observed, even in languages that stand very close to the ideal prototype of Niger-Congo noun class systems (which is uncontroversially the case of Tswana).

A1.3.1. Class is not always overtly indicated in noun morphology

In Bantu, a prefix *n- of class 9 can be reconstructed, but synchronically, in many Bantu languages, no prefix can be isolated in noun forms of class 9; moreover, in Niger-Congo languages, there are not only cases of nouns that invariably show a zero class affix: optional class marking is unfrequent, but not totally unknown.

(5) *Optional deletion of the class prefix of Tswana nouns belonging to classes 5, 7, 8-10, 11 or 14 combined with modifiers*

- a. *lesaka la dikgomo* = *saka la dikgomo*
 lì-sàkà l-á-dí-q^hòmú
 CL5-cattle_kraal CL5-GEN-CL8/10-cattle
 ‘cattle kraal’
- b. *dikgomo tsa kgosi* = *kgomo tsa kgosi*
 dì-q^hòmú ts-á-q^hósí
 CL8/10-cattle CL8/10-GEN-CL9-chief
 ‘the cattle of the chief’

A1.3.2. Two nouns with different agreement properties may show phonologically identical class markers

In such cases descriptions of Niger-Congo languages generally consider that the two nouns belong to distinct classes with homonymous class markers, which is consistent with the definition of noun classes as sets of forms having identical agreement properties.

This situation is illustrated by Bantu classes 1 and 3 (which in Tswana have the same nominal prefix *mo-* [mù]). In the evolution of noun class systems, the homophony of noun affixes may lead to a confusion in their agreement properties, in particular when two or more classes include nouns with zero prefixes. For example in Tswana, there is a considerable degree of hesitation in assigning names of animals to class 9 or to class 1a, which both have a zero noun prefix; assignation to class 9 seems to be etymological, but re-assignation to class 1a is favored by the fact that class 1 typically includes names of human beings, and the protagonists of traditional tales are humanized animals whose names trigger class 1 agreement, typically associated with names of human beings, instead of the agreement they trigger when denoting ordinary animals.

A1.3.3. Two nouns with different class markers may have exactly the same agreement properties

Linguists describing Niger-Congo languages, when faced with such situations, do not always retain the same solution. They sometimes consider that two classes may be distinguished in noun morphology only. However, this solution is logically open to criticism, since the very notion of noun class crucially relies on agreement, and in practice it may lead to a needless proliferation of ‘classes’. Nouns with the same agreement properties but with different class markers are better analyzed as belonging to two sub-classes of the same class. For example, descriptions of Bantu languages traditionally recognize a subclass 1a of class 1. Nouns of class 1a have exactly the same agreement properties as the other nouns of class 1, but nouns of class 1 generally have an overt class marker (Tswana *mo-* [mù]), whereas nouns of class 1a have a zero class marker.

A1.3.4. Morphemes involved in the same agreement pattern are far from being always phonologically uniform

Contrary to what may suggest handbook examples that heavily emphasize the regularity of Niger-Congo agreement systems, morphemes involved in the same agreement pattern (i.e. morphemes that indicate the same class but attach to words of different syntactic status – nouns, noun modifiers, pronouns or verbs) are far from being always phonologically uniform. For example, in Tswana, as illustrated by the examples above, classes 1, 3, 4, 6, 9 and 15-17 involve particularly heterogeneous sets of markers. This situation may have its origin in the conflation of classes originally distinct, resulting in the emergence of a class mixing morphological elements that originally belonged to distinct classes. It is however important to emphasize that, synchronically, agreement rules apply with the same regularity, irrespective of the degree of phonological heterogeneity of the sets of markers they involve.

A1.4. Variations in the scope of class agreement rules

In Tswana, in the construction of noun phrases, all types of noun dependents without any exception express class agreement with the head noun, and verb forms include a subject marker expressing class agreement with the subject NPs. In less typical noun class systems, some types of noun dependents do not agree with their head. In Atlantic languages, this is often the case with genitives. It may also happen that class agreement is limited to the construction of the noun phrase. For example, in Wolof (Atlantic), noun dependents express class agreement with their head, but verbs agree with their subject in person and number only, not in class, and similarly, object pronominal clitics have 3rd person forms that neutralize class distinctions.

A1.5. Semantic aspects of Niger-Congo noun class systems

Elementary presentations have popularized the idea that Niger-Congo classes have very transparent meanings, but things are not so simple. The assignation of individual nouns to genders is clearly neither random, nor predictable on the basis of fairly obvious notions such as ‘animals’, ‘trees’, ‘body parts’, etc., as sometimes suggested.

A1.5.1. Lack of semantic transparency

In Niger-Congo noun class systems, \pm human is the only semantic distinction whose relevance is absolutely obvious. Other semantic features correlate with gender assignment, but in a less obvious way. Regularities are never absolute, but rather of statistical nature, and consequently it is difficult to test the relative validity of possible systems of semantic features and to determine which aspects of the lexical meaning of nouns must be considered as particularly relevant to gender assignation.

A1.5.2. Semantic motivation vs. arbitrariness in the treatment of loanwords

The treatment of borrowed nouns confirms the complexity of the relationship between gender assignment and the lexical meaning of nouns. Borrowed nouns may be integrated to the noun class system on a semantic basis (i.e. they may be assigned to a gender that already includes nouns with which they share some semantic feature),⁸ but very often, gender assignation is determined by other factors, which contributes to weaken the semantic regularities underlying noun classification. In a language with class prefixes, if the initial of a borrowed noun is phonologically similar to some class prefix, it may be reanalyzed as representing the prefix in question, as illustrated by Tswana *bo-rotho* [bùròtʰó] pl. *ma-rotho* ‘bread’, from Dutch *brood*. But in Niger-Congo noun class systems, it is particularly common that most borrowed nouns are assigned to a gender that consequently can be considered as the

⁸ For example, in Tswana, *mo-fine* ‘wine’ (borrowed from Dutch) belongs to gender 3/4; this is clearly neither a phonologically motivated assignation nor a default assignation, and the only possible explanation is that gender 3/4 includes nouns of other psychoactive substances such as *mo-re* ‘medecine’, *mo-tsoko* ‘tobacco’, *mo-tokwane* ‘marijuana’.

default gender for borrowed nouns. Very often, the default gender for borrowed nouns is characterized by the absence of overt class marker in the singular form of nouns.

A1.5.3. Semantic motivation in the ‘derivational’ use of class markers

If inflection is defined as the part of morphology directly involved in syntactic rules, the class prefixes of nouns are unquestionably inflectional. However, noun class systems tend to blur the traditional inflection vs. derivation distinction in the sense that, in addition to number, commonly integrated into noun inflection, class alternations express meanings more commonly expressed by derivational morphology.

The observation of class alternations expressing ‘derivational’ meanings reveals semantic motivations that are not immediately apparent in the assignation of individual non-derived noun stems to genders defined on the basis of the expression of number. Noun stems may be involved in multiple class alternations, with meanings such as *tree vs. fruit, individual vs. collective, concrete vs. abstract, diminutive, augmentative, etc.*

(6) *Two examples of Tswana noun stems whose combination with class prefixes is not limited to the expression of number*

- | | | | |
|----------|------------------|--------------------------------------|-------------------------------|
| a. cl. 1 | <i>mosadi</i> | [mòsádí] | ‘woman’ |
| cl. 2 | <i>basadi</i> | [bàsádí] | ‘women’ |
| cl. 7 | <i>sesadi</i> | [sìsádí] | ‘feminine behavior’ |
| cl. 9 | <i>tshadi</i> | [ts ^h ádí] | ‘group of women’ |
| cl. 11 | <i>losadi</i> | [lòsádí] | ‘group of women’ |
| cl. 14 | <i>bosadi</i> | [bùsádí] | ‘womanhood’ |
| | | | |
| b. cl. 3 | <i>moretlwa</i> | [mòrét ^{wá}] | ‘tree of the sp. moretlwa’ |
| cl. 2 | <i>meretlwa</i> | [mìrét ^{wá}] | ‘trees of the sp. moretlwa’ |
| cl. 9 | <i>thetlwa</i> | [t ^h ét ^{wá}] | ‘fruit of the moretlwa tree’ |
| cl. 10 | <i>dithetlwa</i> | [dìt ^h ét ^{wá}] | ‘fruits of the moretlwa tree’ |
| cl. 11 | <i>loretlwa</i> | [lòrét ^{wá}] | ‘thicket of moretlwa trees’ |

(7) *The relative regularity of the expression of ‘derivational’ meanings by means of class alternations (1 vs. 14 and 3 vs. 9) in Tswana*

- | | | | | | | |
|--------------------|------------------------|-----------------|---|----------------|------------------------------------|------------------|
| a. <i>mosadi</i> | [mòsádí] | ‘woman’ | / | <i>bosadi</i> | [mòsádí] | ‘womanhood’ |
| <i>monna</i> | [mònná] | ‘man’ | / | <i>bonna</i> | [bònná] | ‘manhood’ |
| <i>moloi</i> | [mòlòì] | ‘witch’ | / | <i>boloi</i> | [mòlòì] | ‘witchcraft’ |
| | | | | | | |
| b. <i>moretlwa</i> | [mòrét ^{wá}] | ‘moretlwa tree’ | / | <i>thetlwa</i> | [t ^h ét ^{wá}] | ‘moretlwa fruit’ |
| <i>morula</i> | [mòrúlá] | ‘morula tree’ | / | <i>thula</i> | [t ^h úlá] | ‘morula fruit’ |
| <i>mmilo</i> | [m̀míló] | ‘mmilo tree’ | / | <i>pilo</i> | [p ^h íló] | ‘mmilo fruit’ |

A1.5.4. *Deverbal nouns and adjectives used as nouns*

There are also obvious semantic regularities in the assignation of deverbal nouns or nominalized adjectives to classes. For example in Tswana, adjectives are mainly used as noun modifiers, and in this use their prefix is determined by agreement rules that operate without taking into account the semantic nature of the noun they modify; but in some classes (not all), adjective stems may also be independently used as nouns with a type of meaning determined by the class prefix, and the meanings expressed are consistent with those expressed by class alternations with noun stems.

(8) *Class alternations with noun stems and class assignation of adjective stems used as nouns in Tswana*

- | | | | | | | |
|----|---------------|----------|---|---------------|----------|-------------------------------------|
| a. | <i>mosadi</i> | [mòsádí] | / | <i>basadi</i> | [bàsádí] | ‘woman / women’ |
| | <i>bosadi</i> | [bòsádí] | | | (cl. 14) | ‘womanhood’ |
| | <i>sesadi</i> | [sìsádí] | | | (cl. 7) | ‘feminine behavior’ |
| b. | -ntle | [ñtłè] | | | | ‘good, beautiful’ (adjectival stem) |
| | <i>bontle</i> | [bòñtłè] | | | (cl. 14) | ‘goodness, beauty’ |
| | <i>sentle</i> | [sìñtłè] | | | (cl. 7) | ‘well’ |

A1.5.5. *The ‘absolute’ use of modifiers*

More generally, it is not uncommon that words marked for class and typically used to modify nouns or to anaphorically refer to nouns also have ‘absolute’ uses in which they are neither syntactically nor anaphorically linked to any noun. In such uses, the class prefix itself conveys a meaning, instead of expressing the class of a possible antecedent.

(9) *Examples of meanings conveyed by class prefixes in the ‘absolute’ use of modifiers in Tswana: 1/2 → persons, 7/8-10 → things*

- | | | | | | | | |
|----|-----------------|---------------------|-------------|---|------------------|------------------------|--------------------|
| a. | <i>mongwe</i> | mù-ŋwì | ‘someone’ | / | <i>sengwe</i> | sì-ŋwì | ‘something’ |
| | | CL1-one | | | | CL7-one | |
| b. | <i>ope</i> | ú-pé | ‘nobody’ | / | <i>sepe</i> | ú-pé | ‘nothing’ |
| | | CL1-not_any | | | | CL7-not_any | |
| c. | <i>ba motse</i> | b-á-mú-tsí | ‘villagers’ | / | <i>tsa motse</i> | ts-á-mú-tsí | ‘things concerning |
| | | CL2-GEN-CL3-village | | | | CL8/10-GEN-CL3-village | the village’ |

This absolute use of modifiers with prefixes carrying general meanings such as ‘person’, ‘thing’, ‘place’ or ‘manner’ can be viewed as a borderline case of a discourse phenomenon by which head nouns are commonly dropped in contexts in which they are relatively easy to retrieve, even if there is no anaphoric relation in the usual sense of this term. For example, in Tswana, class markers of class 5 can be interpreted as referring to *letsatsi* ‘day’ even in the absence of any syntactic or anaphoric link, provided this default interpretation is semantically consistent with

the rest of the sentence and is not in competition with a possible anaphorical link with some other noun of class 5 that would provide a more consistent reading: instead of *ka letsatsi le le latelang* lit. ‘on the day that follows’, it is possible to say simply *ka le le latelang*, lit. ‘on the one (cl.5) that follows’, and if the context does not suggest an anaphorical link with some other class 5 noun, the hearer will immediately interpret this as ‘on the following day’.

A1.5.6. ‘Pronominal classes’

In Niger-Congo noun class systems, as mentioned in the previous section, class markers included in pronominal forms are not always triggered by agreement in a syntactic configuration or reference to a entity or notion already mentioned or present in the speech situation. They may also directly encode generic meanings such as ‘person’, ‘thing’, ‘manner’, ‘time’ or ‘place’.

Moreover, in addition to class markers regularly involved in the formation of words that in some way or other agree with nouns, pronoun formation may involve morphemes whose distribution is only part of the distribution of regular class markers. Such morphemes are found in forms whose morphological structure is identical to that of ordinary pronouns, they occupy morphological slots normally occupied by class markers expressing agreement with nouns, but they never express agreement with nouns, and are used exclusively to express generic notions such as ‘place’, ‘manner’, or ‘time’.

Jóola-Banjál (Atlantic) has two such morphemes, *d(i)*- ‘place’ and *n(i)*- ‘time’. For example, *d-áno-d-an* ‘anywhere’ and *n-áno-n-an* ‘at any time’ are built according to the same pattern *CL-áno-CL-an* as *f-áno-f-an* ‘any (cl. 7)’ or *y-áno-y-an* ‘any (cl. 3)’. But *f-áno-f-an* and *y-án-o-y-an* can be used as noun dependents in constructions characterized by class agreement (*fu-mango f-áno-f-an* ‘any mango’, *e-kulol y-áno-y-an* ‘any chicken’, whereas *d-áno-d-an* and *n-áno-n-an* cannot be used as noun modifiers.

A plausible explanation is that such morphemes originally marked agreement with nouns expressing the meanings in question, and that the nouns with which they marked agreement subsequently disappeared, or were transferred to other classes.

A1.5.7. *More on the relevance of the human vs. non-human distinction in Niger-Congo noun class systems*

In Niger-Congo languages, class agreement in conformity with the class prefixes of nouns tends to be replaced by a semantic agreement rule that neutralizes class distinctions and takes into consideration the *human vs. non-human* or *animate vs. inanimate* distinction only:

- when a pronoun and its antecedent do not belong to the same clause;
- when agreement involves coordinated NPs.

In some languages, semantic agreement neutralizing class distinctions extends to human or animate nouns in other contexts. For example, in Jóola-Banjál (Atlantic), human nouns trigger agreement in conformity with their class prefix in combination with some types of noun modifiers only. Other types of modifiers invariably show

gender 1/2 agreement if they modify a human noun, irrespective of its class prefix, and the same applies to verbs agreeing with a human subject.

A radical reorganization of the noun class system along these lines is well attested among Gur and Kwa languages, and is found in some Bantu languages too. In the languages in question, the former division of nouns into classes may still be apparent in noun morphology, but the relationship between noun prefixes and the agreement properties of nouns has been completely lost, a reduced number of agreement patterns (typically four) has been retained, and agreement is directly triggered by number and animacy.

A1.6. Conclusion

All the available evidence points to a very ancient origin of Niger-Congo noun class systems. The classification system reconstructed for Proto-Bantu is very similar, in all respects, to the systems of many modern Bantu languages. As for Niger-Congo, there is so far no real reconstruction of a proto-language at this level, but the other branches of Niger-Congo do not seem to provide any evidence supporting the reconstruction of a less grammaticalized noun class system at Niger-Congo level. Noun class systems that can be viewed as ‘incomplete’ by comparison with the Bantu prototype are very common in various branches of Niger-Congo, but historically, they are not emerging class systems, but rather the result of the disintegration of former systems of the Bantu or Atlantic type.

In particular, it must be emphasized that the comparison of Niger-Congo languages has so far revealed no concrete evidence pointing to a possible lexical origin of class markers. More generally, Niger Congo noun class systems do not seem to have preserved any trace of stages of their evolution in which we can imagine that they were characterized by a lesser degree of grammaticalization. In other words, in the present state of our knowledge, hypotheses about the genesis of Niger Congo noun class systems are condemned to remain purely speculative.

By contrast, Niger-Congo languages provide abundant illustration of processes that noun classification systems at an advanced stage of grammaticalization may undergo, in particular:

- renewal of class morphology by agglutination of former determiners to nouns,
- changes in the gender assignment of individual lexemes, either formally or semantically motivated,
- conflation of originally distinct genders,
- disintegration of systems of agreement rules,
- evolution of highly grammaticalized noun class systems with a relatively high number of genders towards two gender systems transparently based on the \pm human or \pm animate distinction.

Appendix 2. Head marking in the genitival construction of two West African languages

Genitival constructions involving a special form of the head noun (‘construct form’) are common among Nilotic languages, but are sporadically found in other parts of Africa too.

Hausa (Afroasiatic, Chadic) has a construct form of nouns characterized by a suffix *-n* (singular masculine or plural) or *-ř* (singular feminine), commonly called ‘genitive linker’. This suffix occurs when the noun fulfills the role of head in the genitive construction – ex. (1a) & (1c). It must also be used when the noun takes a possessive suffix other than 1st person singular – ex. (1e). It results from the cliticization of a pronoun *na* / *ta* co-referent with the head noun in the synonymous construction illustrated by ex. (1b) & (1d).

- (1) a. *kàre-n Daudà* (cf. *kàree* ‘dog’)
 dog-H:SG:M Dauda
 ‘Dauda’s dog’
- b. *kàree na Daudà*
 dog that_one_(SG:M)_of Dauda
 ‘Dauda’s dog’
- c. *saaniya-ř Daudà* (cf. *saaniyaa* ‘cow’)
 cow-H:SG:F Dauda
 ‘Dauda’s cow’
- d. *saaniyaa ta Daudà*
 cow that_one_(SG:F)_of Dauda
 ‘Dauda’s cow’
- e. *kàree-naa* ‘my dog’ *saaniyaa-taa* ‘my cow’
 kàre-n-kà ‘your(SG:M) dog’ *saaniya-r-kà* ‘your(SG:M) cow’
 kàre-n-kì ‘your(SG:F) dog’ *saaniya-ř-kì* ‘your(SG:F) cow’
 kàre-n-sà ‘his dog’ *saaniya-ř-sà* ‘his cow’
 kàre-n-tà ‘her dog’ *saaniya-ř-tà* ‘her cow’
 kàre-n-mù ‘our dog’ *saaniya-ř-mù* ‘our cow’
 kàre-n-kù ‘your(PL) dog’ *saaniya-ř-kù* ‘your(PL) cow’
 kàre-n-sù ‘their dog’ *saaniya-ř-sù* ‘their cow’

A difficulty in the analysis of *-n* ~ *-ř* as the mark of a construct form of Hausa nouns is however that the same suffix characterizes attributive adjectives preceding nouns in the construction illustrated by *fari-n kàree* ‘white dog’ / *fara-ř saaniyaa* ‘white cow’, where *fari* / *fara* are the masculine and feminine forms of the adjective ‘white’. In this construction, *-n* ~ *-ř* cannot be recognized as an instance of head marking. A possible solution is to consider that, when attributive adjectives precede nouns, they take an additional gender agreement mark homonymous with the suffix of the construct form of nouns (and certainly cognate with it in a historical perspective).

In Wolof (Niger-Congo, Atlantic), a construct form of nouns characterized by the suffix *-u* (sg.) / *-i* (pl.) is used exclusively with nouns combined with another noun in the role of genitival dependent. It occurs with no other dependent, and, contrary to Semitic construct forms, it does not occur with possessives either.

The construct form of Wolof nouns shares with Semitic construct forms a constraint of strict contiguity with the dependent noun. This means that other dependents of the head noun in the construct form must follow the genitival

dependent, and that, if the dependent noun itself has dependents that must precede it, they must be placed to the left of the head noun in the construct form, as illustrated by ex. (2).

- (2) a. *fas wu ñuul*
horse CL_W.LNK be_black
'black horse'
- b. *suma nijaay*
1SG maternal uncle
'my uncle'
- c. *suma fas-u nijaay wu ñuul*
1SG horse-H maternal uncle CL_W.LNK be_black
'the black horse of my uncle' (lit. 'my horse of uncle black')
- d. **fas-u suma nijaay*
horse-H 1SG maternal uncle