Towards a Typological Framework for Niger-Congo Noun Classes: Challenges and Criteria

Socroun, Clémentine

Master's Thesis
MA Linguistics-Linguistics
Leiden University

***Towards a Typological Framework for Niger-Congo Noun Classes:
Challenges and Criteria***

Clémentine Socroun

Supervisor: Dr. Christian J. Rapold
Second reader: Dr. Erik van Gijn

Leiden, The Netherlands
June 30, 2021

Abstract

This thesis aims at filling a gap in the typological analysis of Niger-Congo noun classes and initiate the creation of a framework applicable for every Niger-Congo noun class system. It answers the research question "What are the criteria and challenges to creating a typological framework for Niger-Congo noun classes ?" Based on existing literature, this thesis reviewed the definitions of noun classes and existing frameworks in most of Niger-Congo language families. It also identifies the challenges specific to this type of project and develops an approach to overcome the difficulties by using the resources of technology, typically Excel, to build a database, sufficiently general to take into account all Niger-Congo noun class systems and flexible enough to adapt to all specificities. Such a framework also aims at creating a database in order to create statistical analysis.

Table of Contents

List of Abbreviations

AGR : Agreement class

NCA : Noun Class Agreement

NCM : Noun Class Morpheme/Marker

NF : Noun Form class

NP : Nominal Phrase

VP : Verbal Phrase

Introduction

The languages spoken on the African continent are divided into four major phyla: Afro-Asiatic languages (north), Nilo-Saharan languages (south Sahara), Khoisan languages (southern Africa) and Niger-Congo languages (equatorial and central Africa). Our study focuses on the family of Niger-Congo languages, which includes more than 1,500 languages. This family alone accounts for 21% of the world's languages. The Niger-Congo languages constitute the most extensive of African language families, both in terms of geographical distribution and number of speakers. These languages are spoken by about 500 million speakers, i.e. 7% of the world population and 85% of the African population. Half of Niger-Congo languages are spoken by less than 29,000 speakers.

The main characteristic of Niger-Congo languages is the use of a system of noun classes. Creissels considers that "one finds in almost all the language families constituting the Niger-Congo phylum a more or less important proportion of languages attesting to the particular type of nominal classification known as noun classes, or at least vestiges of a system of this type. The most typical languages from this point of view are the Bantu languages and the Atlantic languages"(Creissels, 2001, p. 157)[1]. Not all Niger-Congo languages have a noun class system, such as Mande, Ijoid, Dogon and Ubangi, which are therefore excluded from the scope of our study, which will only focus on Niger-Congo languages with a system, even a partial one, of noun classes.

For the purposes of this thesis, we have adopted the definition of Robert Hepburn-Gray (2018), who considers a noun class to be: "a unique pairing of noun form and agreement class exhibited in the behavior of a noun". The description of noun class systems in Niger-Congo languages differs from one family to another, from one language to another and even from one linguist to another. Indeed, the linguistic phenomenon of noun classes, which has been studied and analyzed extensively, divides the community of linguists who oppose each other with different approaches. There is no simple and unique way of understanding noun classes. Thus, it is difficult to establish generalities, common to all Niger-Congo noun classes, and to have a global vision. Indeed, there is no typological framework,

---

[1] « On trouve dans presque toutes les familles de langues constituant le phylum Niger-Congo une proportion plus ou moins importante de langues attestant le type particulier de classification nominale connu sous le nom de « classes nominales », ou du moins des vestiges d'un système de ce type. Les langues les plus typiques de ce point de vue sont les langues bantoues et les langues atlantiques. » (Creissels, 2001, p. 157)

common to all Niger-Congo languages, which would allow comparisons between the noun classes of different languages on a common ground, to identify similarities and differences, and which would allow linguists to speak the same language and to communicate better among themselves and with other disciplines on this theme. Creissels himself recognizes the urgency of setting up such a common typological framework (2015, p.4): "The current situation regarding classes in Niger-Congo languages, in particular in Atlantic languages, has as a corollary that arbitrary numbering by this or that author constitutes a real obstacle to the discussion of classes. Indeed, there is for the moment no possibility of extending the principle that governs the numbering of Bantu classes to the whole of Niger-Congo in a coherent and exhaustive manner, since this would imply systematically numbering the classes of Niger-Congo languages according to their relation to a reconstructed Niger-Congo system[2]."

If such a typological framework meets the genuine expectations of the community of Africanists, typologists, and students of linguistics, why doesn't it exist? Undoubtedly, because there are obstacles and specificities to its implementation that could not easily be overcome. This is what we will study by answering the Research Question: "What are the criteria and challenges to creating a typological framework for Niger-Congo noun classes? "

This research question generates four distinct questions that correspond to the four chapters of this dissertation and will allow us to understand how to set up such a typological framework. In the first chapter we will attempt to define what noun classes are in Niger-Congo languages and to draw a global definition. In the second chapter we will observe existing frameworks and understand what gap a common framework would fill for the description and analysis of Niger-Congo noun classes. In the third chapter we will analyze the challenges that must be overcome to create a typological framework. Finally, in the fourth and last chapter, we will define the criteria and features necessary for the creation of a typological framework for Niger-Congo languages, and we will present a matrix that could serve as a basis for the creation of this general framework.

To successfully answer the Research Question and the questions that arise from it, we have put in place a methodology that relies largely on research in literature on Niger-Congo noun classes, together with Excel software. Thus, we developed a

---

[2] « *La situation actuelle en ce qui concerne les classes dans les langues Niger-Congo, et en particulier dans les langues atlantiques, a pour corollaire qu'une numérotation arbitraire par tel ou tel auteur constitue un véritable frein à la discussion sur les classes. En effet, il n'y a pour l'instant aucune possibilité d'étendre de manière cohérente et exhaustive à l'ensemble Niger-Congo le principe qui préside à la numérotation des classes bantoues, car cela supposerait de numéroter systématiquement les classes des langues Niger-Congo selon leur relation à un système Niger-Congo reconstruit.* »

research plan to allow us to bring new elements to support our argument. We then collected data empirically, pooled them together and analyzed them. With all this information, we were able to write the following dissertation.

CHAPTER 1: Noun Classes Definition(s)

Introduction

Noun classes constitute the main characteristic of Niger-Congo languages, which include almost all the languages spoken in Africa south of the Sahara, i.e. more than 1,500 languages and about 500 million speakers. The system of noun classes is found in most of the branches of Niger-Congo languages. Still, it remains unknown in some languages such as Ijoid, Mande and Dogon, where there is no trace of them.

The first chapter of this paper deals mainly with definitions and historical context. This chapter describes the historical context that is fundamental to properly understand Niger-Congo noun classes and their descriptions. By briefly recalling the link between the progress of colonization from the 15th century onwards and the development of the study of south of the Sahara languages, it explains why and how noun classes have been studied mainly for Bantu languages. It also explains why it is only recently, in the 20th century until today, that noun classes have been studied for all Niger-Congo languages.

In addition, this first chapter attempts to establish a definition of the phenomenon of noun classes in Niger-Congo languages. It inventories the different approaches that have marked out research at different times and it studies the main points in order to come up with a definition that is both sufficiently complete and simple to serve as a basis for the rest of our study.

1. Historical Context

The first instances of writings by Portuguese priests that can undoubtedly be associated with Niger-Congo languages can be traced back to 1506. Indeed, the Catholic priests who left Portugal to evangelize the regions of Congo and Lower-Congo had to learn some Kikongo or Kongo language to communicate with local population. "The linguistic production of a few Franciscans, Augustinians, Capuchins, Dominicans, and/or Jesuits, working under the aegis of the Portuguese Crown, who […] compiled the first grammars, word lists, glossaries, and dictionaries of the indigenous languages with which they worked and interacted on a daily basis" (Levi, 2009, p.391). It was in 1624 that the first literature partly written in a Niger-Congo language was published. A catechism, the "Doutrina Christaã", written

by three Jesuits, under the direction of Mateus Cardoso, in Portuguese with an interlinear translation in Kongo.

The Kongo kingdom was located in central Africa, at the mouth of the river Zaire. It was explored by Portuguese colonialists as of the 15th century and later assimilated to the colony of Angola. This historical background accounts for why Bantu languages were the first African languages studied and documented by Western researchers.

In 1650, Giacinto Brusciotto, an Italian missionary, wrote the first multilingual dictionary of Kongo -lost since then- which included explanations in Portuguese, Latin and Italian. Brusciotto published the first grammar of Kikongo in 1659, entitled "Regulae quoedum pro difficilimi congensium idiomatis faciliori captu ad grammaticae normam redactoe". Translated by Doke in 1959 (Maho, 1999) as: "Some rules for the easier understanding of the most difficult idiom of the people of the Congo, brought into the form of a grammar".

It is in this book that the features of noun classes were described for the first time without naming them as such yet.

Brusciotto's classification was mainly based on morphology and syntax because of the importance given to agreement patterns. As Maho puts it, citing Doke (1959), "Bantu noun classes were first described by Brusciotto who "coined the concept of "principiationes" to indicate the categories into which nouns fell, according to their first element but he focused on the agreement class pattern more than on the noun class marker" (Maho,1999, p.13). Brusciotto pointed out that all nouns have an element that marks the way they combine with other words in the sentence (Bendor-Samuel, 2006). He thus identified 18 "principiatones" / prefixes that determine the rules of agreement. This first analysis of the Kongo language was still influenced by a vision dependent on European languages (especially declensions). Still, according to Bendor-Samuel "Brusciotto's work is praised for its accurate understanding of the noun and verbal systems of Kongo, despite the lack of analogous systems in Latin or any other previously studied grammars".

If noun classes have been mainly described for the Bantu languages, it is because these languages have been the subject of the first and most thorough studies, which have led to the reconstruction of proto-Bantu. The foundations laid down by Brusciotto are still valid today, even if they have been deepened since then. "African linguistics developed in the 19th century, when European missionaries studied African languages in some detail. According to some authors, salient typological properties of African languages, such as complex tone structures and noun class systems, were identified and described" (Heine, 2006). It was at this time that the

term "noun classes" appeared in "A comparative study of South African languages" (Bleek, 1862), in which classes were analyzed by separating prefixes and their concords.

According to McCormack (2007), William Marsden (1818) described the contours of the Bantu family in his comparative study of Bantu and semi-Bantu languages. He initiated the work of the German linguist, Wilhelm H. J. Bleek, who "laid down most of the fundamental principles still accepted by modern Bantuists"[3] (Alexandre, 1959, p.298). We owe to him the term "Bantu" in reference to the Zulu word aba-ntu meaning people/men" to cover a group of African languages that includes about 500 languages spoken in about 20 countries in the southern part of Africa. Bleek also laid the basis for a classification system in which the noun classes of the Bantu languages are numbered. This system initially comprised 16 classes of nouns marked by prefixes.

It is around the same period that the German missionary Sigismund W. Koelle published, in 1854, a study entitled "Polyglotta Africana", in which he compared 280 words from 200 Western African languages and dialects (120 languages according to today's classification). Koelle did not go further in his analysis, but his book attests to the existence of noun classes in Atlantic languages, not only in Bantu languages. From this perspective, Koelle is the first author to establish a relation between Bantu languages and North-West Atlantic ones through noun classes, which are distinguished by prefix changes or an initial inflection as referred to by Koelle.

The German linguist Carl Meinhof went further. He presented the first outline of reconstructed proto-Bantu, by comparing different class systems from several Bantu languages. He followed Bleek's numbering and added some classes to classify Bantu languages into 21 noun classes. While there is no language that expresses all of them, most Bantu languages have at least 10. For example, Swahili has 15, Sesotho 18, and Luganda 19. The list begins with the most obvious and widely attested classes, such as the odd numbers singular classes and the even numbers plurals, generally plurals of the immediately preceding singulars. After listing the classes that commonly appear in singular-plural pairs, the numbering continues for the classes that are not participating in the SG-PL distinction. Since Meinhof's day, a few classes have been added to the list that have been found in only a few languages, and some refinement in reconstruction have been achieved. Clement Doke, referring to kinship, added classes 1a and 2a, Desmond T. Cole added 2b and 8x, and Welmers added class 6a, referring to liquid masses. Msaka considers that "despite such revisions, the Bleek-Meinhof framework has remained the standard" (Msaka, 2019) especially for Bantu languages.

---

[3] "W. Bleek […] pose la plupart des principes fondamentaux encore admis par les bantouistes modernes."

It was during this period that researchers also established that noun classes are not the result of an evolution but rather the survival of a proto-historical state. Most linguists agree that the Proto-Niger-Congo had a noun class system, although not all Niger-Congo languages retained it. Many languages show partial retention with, for example, a reduced system of a small number of classes or the survival of traces of noun class system whose concordant characteristics have been lost. As Creissels puts it, "the system reconstructed for proto-Bantu is very similar to what is found in many of today's Bantu languages. As for Niger-Congo, there is no real reconstruction of a proto language yet, but specialists agree that there are no indications that it would support the reconstruction of a less grammaticalized classification system. "Incomplete noun class systems" in relation to the Bantu prototype are very common in various branches of Niger-Congo, but all indications show that they should not be interpreted as an emerging class systems, but rather as the results of the disintegration of older systems close to the Bantu prototype"[4] (Creissels, 2001).

The system of noun classes is found, in one form or another, in most of the branches of the Niger-Congo family. Still, it remains unknown in some languages such as Ijoid, which are known for their subject-object-verb structure, which is unusual in Niger-Congo languages, a phenomenon that is only shared by distant branches such as Mande and Dogon. In these languages, there is no trace of noun classes, characteristic of Niger-Congo languages.

Subsequently, de Wolf (1971), who participated in the reconstruction of Proto-Benue-Congo[5], deepened Bleek's classification by establishing a new classification table in his book "The noun class system of Proto-Benue-Congo". He states "In the wider Benue-Congo context, the most significant work is de Wolf (1971), which reconstructs a noun class system for Proto-Benue-Congo and remains the most detailed study on the topic available today" (Good & Creissels, 2018).

---

[4] "Le système reconstruit pour le proto-bantou est très semblable à ce qu'on trouve dans beaucoup de langues bantoues actuelles. Quant au Niger-Congo, il n'y a pour l'instant aucune réelle reconstruction d'une proto-langue à ce niveau, mais tous les spécialistes s'accordent pour dire qu'on ne trouve aucune indication qui irait dans le sens de la reconstruction d'un système de classification moins grammaticalisé.9 Les systèmes de classes nominales « incomplets » par rapport au prototype bantou sont très communs dans diverses branches du Niger-Congo, mais tout indique qu'ils ne sont pas à interpréter comme des systèmes de classe émergents, mais plutôt comme le résultat de la désintégration de systèmes plus anciens proches du prototype bantou"

[5] Benue–Congo is one of the largest subdivisions of the Niger–Congo language family, both in number of languages, of which Ethnologue counts 976, and in speakers, numbering around 350 million.

Table 1: Overview of de Wolf's (1971) Proto-Benue-Congo noun class system

LABEL	PFX	CONC	PFX	CONC	EXAMPLE
1/2	*ù-, *ò-	*gwu-, *à-	*bà-	*ba-	*-lume 'man'
5/6	*li-	*zí- (?)	*à-	*ga-, *a-	*-tama 'cheek'
7/8	*ki-, *ke-	*ki-	*bì-, *bè-	*bi-	*-kupe 'bone'
9/10	*e-, *i-	*zì-	*í-	*í-, *zi- (?)	*-name 'animal'
3	*ú-	*u-, *gu-	(plural 10 or 13)		*-sene 'back'
11	*lu-	*lu-	(plural 10)		*-z(w)ana 'sun'
12	*kà-	*ka-	(plural 10 or 13)		*-kama 'monkey'
13	(singular 3 or 12)		*ti-	*ti-	*-kwon 'tree'
14	*bù-	*bu-	(plural 6 or 10)		*-su 'face'; *-bogo 'fear'
15	*ku-	*ku-	(plural 6 or 10)		*-tuŋi 'ear'
6a	*mà-, *nà-	*ma-, *nà-	(unpaired)		*-luŋ 'blood'

Figure 1: Overview of de Wolf's (1971) Proto-Benue-Congo Noun Class System

2. Different Descriptions and Definitions

It appears that there are two major definitions of how noun classes are approached by linguists. One is based on semantics, with nouns being grouped according to their meaning (animate/inanimate, human/non-human, …). The other one is based on grammatical rules, with nouns being grouped according to the concordance agreement they govern.

In addition to these major common features, the system of noun classes in Niger-Congo languages is described with a variable level of detail according to authors and periods (cf. Chapter 2). Most researchers believe that noun classes "constitute an obligatory grammatical system, where each noun chooses one from a small number of possibilities" (Dixon, 1986, p.105) and that they are defined by prefixes and/or suffixes that allow nouns to be grouped. "The tradition is to reserve the appellation "gender" for systems based on the semantic trait of "sex/M-F", taking Indo-European languages as the prototype cases, while the appellation of "noun classes" applies to systems with a larger number of classes (considered exotic from an Indo-European language-centered point of view), with the systems of Bantu languages as the prototype"[6] (Grinevald, 1999, p.103).

Some authors stipulate that noun classes represent a set of radicals grouped in relation to a common affix. Hockett (1958, p231) notes that "genders are classes of nouns that are reflected in the behavior of associated words" and that the notions of class and gender are hardly distinct. Creissels is joined by Dixon (1986), and

[6] "La tradition est de réserver l'appellation de "genre" pour les systèmes qui reposent sur le trait sémantique de "sexe/M-F", en prenant les langues indo-européennescomme les cas prototypiques, tandis que l'appellation de "classes nominales" s'applique aux systèmes à plus grand nombre de classes (considérées comme exotiques d'un point de vue centré sur les langues indo-européennes), avec les systèmes des langues bantoues comme prototype."

Corbett (1991), and Grinevald (1999) who consider that gender systems and noun class systems are one in Niger Congo languages. Moreover, Grinevald notes that "gender and noun class systems are treated in recent literature as one fundamental type of noun classification"[7] (Grinevald, 1999, p.103).

For others, the notions of noun classes and genders are considered along paradigmatic and syntagmatic axes. "One calls noun class the set of nouns marked in a similar way and governing the same facts of agreement and substitution"[8] (Bole-Richard, 1983, p.53) or the class groups together "all the nouns that lead to the same choices when the determinants giving rise to agreement are added"[9] (Creissels, 1991, p.82). According to this approach, noun class and gender are distinct. Class stands for an affix with classificatory power. Each affix giving rise to concordance agreement is considered a noun class marker.

As discussed by Corbett (2006), 'alliterative agreement' can be understood in two different ways: "this term may refer to "a characterization of morphological exponence," in systems in which agreement controllers have an inflectional marker correlated to their behavior in the agreement system, and phonologically identical segments are used as agreement markers on agreement targets. In this sense of 'alliterative agreement', Niger-Congo systems of class agreement (but also many Indo-European systems of gender-number-case agreement) can be characterized as partially alliterative, since they involve both class agreement markers phonologically identical to the corresponding class membership markers found in noun forms, and class agreement markers phonologically distinct from the corresponding class membership markers of nouns." As rightly pointed out by Corbett, this characterization of agreement systems as ±alliterative is not an "all or nothing" classification. Moreover, it is worth emphasizing that a thorough description of Niger-Congo class agreement systems often leads to the conclusion that they are in fact much less alliterative than they may look at first sight, because class agreement markers that are roughly like the corresponding class membership markers of nouns often differ from them in an unpredictable way in details such as vowel quality or tone.

For example, in Swahili:
> *In the sentence wa-tu wa- the wa-mefika (composed of a noun, a demonstrative, and a verb, meaning "these people have arrived"), concordant*

---

[7] "les systèmes de genre et de classes nominales sont traités dans la littérature récente comme un seul type fondamental de classification nominale"

[8] "On appelle classe nominale l'ensemble des noms marqués de façon semblable et gouvernant les mêmes faits d'accord et de substitution"

[9] "Tous les substantifs qui entraînent les mêmes choix lors de l'adjonction des déterminants donnant lieu à accord"

elements link the three parts of the sentence by the prefix wa- (Encyclopedia Britannica-Widespread Characteristics of Niger-Congo Lauguages).

Noun class markers (NCM) can appear as prefixes, suffixes, or circumfixes, for example:

NCM prefixing in Baïnounk Gubeëher

bu-rul	i-rul	
CL.bu-mouth	CL.i-mouth	
'mouth'	'mouths'	(Cobbinah 2013:179)

NCM suffixing in Supyire

ba-ga	ba-ya	
house-5	house-6	
'house'	'houses'	(Carlson 1999:142ff)

NCM circumfixing in Ditammari

dī-dù-rì	
5-stick-5	
'stick'	(Reineke 2012:QQ)

It seems that semantic considerations determined initially the affixes that marked a particular noun class, but semantic categories may have faded or changed, and researchers have not discovered the « new » semantics yet. It seems that meaning should always be considered as an indicator of class belonging, but not as much as grammatical evidence. A noun belongs to a given class or gender because of the characteristics of its referent, such as sex, animacy, form, but these designations are often conventional.

3. Characteristics and Theoretical Basis

We note that it is difficult to find a homogeneous, global, and single definition of Niger-Congo noun classes. Indeed, researchers support definitions of their own, resulting from their studies and observations, or don't define their use of noun classes at all. It is difficult to rally them all behind a single definition and it's even harder when it comes to defining the characteristics of noun classes of Niger-Congo languages. Here again, opinions are different. However, researchers are almost unanimous on one point: the system of noun classes is the main feature that characterizes Niger-Congo languages. As Gutman and Avanzati explain: "The system of noun classes is probably the characteristic most widely found in Niger-Congo languages (though Mande does not have them)" (Gutman & Avanzati, 2013). Creissels, for his part, goes even further: "We find in almost all language families constituting the Niger-Congo phylum (which includes the majority of languages

spoken in sub-Saharian Africa) a more or less significant proportion of languages attesting to the particular type of noun classification known as "noun classes", or at least vestiges of such a system. Under this respect, the most typical languages are the Bantu and the Atlantic languages" (Creissels, 2001).

Gutman and Avanzati consider that "nouns are grouped in different classes, marked by prefixes, suffixes or both. All members of a given class share the same affix. The number of noun classes varies greatly from language to language" (Gutman & Avanzati, 2013). In the Atlantic branch, for example, the number of noun classes ranges from 3 to nearly 40. In the Gur branch, there are 11 classes. In the Bantu languages, 12 to 15 noun classes are frequently found (Encyclopedia Britannica-Widespread Characteristics of Niger-Congo Languages).

Corbett (1991) suggests that the use of the labels "class" or "gender" is mostly a matter of tradition. Between African noun classes and European genders, the difference lies essentially in the way these systems are described, not in something deeper.

Indeed, the classification of nouns into genders is based on lexemes, which means that two noun forms considered as the singular and the plural of the same lexeme count in such a classification for a single unit (i.e. the singular and plural forms belonging to the same gender). On the other hand, when speaking of noun classes in Niger-Congo languages, one generally refers to a classification of nouns according to their agreement pattern. The singular and the plural of the same lexeme are counted as two units (the sg and pl beloning to different noun classes). For example, in Tswana, *mo-nna* 'man' (cl.1), *ba-nna* 'men' (cl.2), *le-kau* 'boy' (cl.5) and *ma-kau* 'boys' (cl.6) will be considered as belonging to four different classes. In the nomenclature of Bantu noun classes, the classes to which these four noun forms belong are labeled 1, 2, 5 and 6, respectively. But one could also describe it as a classification in which *monna / banna*, considered as a unit, would be said to belong to gender A and *lekau / makau* to gender B.

In a noun class system, nouns are generally marked with an affix that can, in some instances, be zero. Generally, one affix indicates a singular form, and another one a plural form. Affixes can be prefixes or suffixes, and very rarely infixes, as in Birom, or circumfixes (Nambima, 2017). Their number varies from one language to another.

Most of noun class systems are accompanied by a concordance system, i.e. other elements of the sentence such as determiners, adjectives or numbers. Often verbs are also marked by an affix determined according to the noun class of the subject. If it is an alliterative agreement, the same syllable that marks the noun (or a form that is phonetically close) is often repeated with other elements. Nevertheless, the noun

class agreement form can be different from the noun class prefix that affects the noun, in non-alliterative system. As discussed by Corbett (2006), 'alliterative agreement' can be understood in two different ways: this term may refer to "a characterization of morphological exponence," in systems in which agreement controllers have an inflectional marker correlated to their behavior in the agreement system, and phonologically identical segments are used as agreement markers on agreement targets. In this sense of "alliterative agreement," Niger-Congo systems of class agreement (but also many Indo-European systems of gender-number-case agreement) can be characterized as partially alliterative, since they involve both class agreement markers phonologically identical to the corresponding class membership markers found in noun forms, and class agreement markers phonologically distinct from the corresponding class membership markers of nouns. As rightly pointed out by Corbett (2006), this characterization of agreement systems as ±alliterative is not an "all or nothing" classification. Moreover, it is worth emphasizing that a thorough description of Niger-Congo class agreement systems often leads to the conclusion that they are in fact much less alliterative than they may look at first sight, because class agreement markers that are roughly similar to the corresponding class membership markers of nouns often differ from them in an unpredictable way in details such as vowel quality or tone

It may seem that agreement patterns in Niger-Congo languages are nothing else than simple repetition of the prefix, to the point that one could sometimes speak of an "alliterative" agreement. In fact, it would be a complete misunderstanding, imagining that each class always corresponds to a unique phonological mark that is repeated on all the words. The marks of belonging to a given class vary according to the nature of the noun class. Class agreement pattern in Niger-Congo languages is in fact no more or less "alliterative" than gender agreement pattern in Indo-European languages. A compelling example of a non-alliterative system is the Keerak noun class system. Each marker depending on the concordial element has a different form (Segerer, 2015, p. 121):

Tableau 5. *Synthèse sur les schèmes d'accord*

NOM	PRO	DET	POS	CONN	DEM1	DEM2	DEM3	NUM	ADJ	quel	autre	SUJ1
A	ɔɔ	aw	Ø	ala	ʊmɛ	umu	ʊmma		a	ay	aahʊ	na
JA₂	jɔ	aw		ala ~ jala	ʊmɛ	umu	ʊmma		a	ʊmay ~ ʊjay	aahʊ	na ~ jɪ
BA₁, BU₁	bɔ	ab	b	bala	ʊbɛ	ubu	ʊmba	b	ba	ʊbay	baabʊ	bʊ
BA₂, BU₂	bɔ	ab	b	bala	ʊbɛ	ubu	ʊmba	bʊ	ba	ʊbay	baabʊ	bʊ
E₁	yɔ	ay	y	yala	ʊyɛ	uyu	ʊyya	y	ya	ʊyay	yaayʊ	ɛ
E₂		ay		yala	ʊkɛ	uku	ʊnka	kʊ	ka	ʊyay ~ ʊkay	kaakʊ	kʊ
HA, HU	hɔ	ah	h	hala	ʊhɛ	uhu	ʊŋha	h	ha	ʊhay	haahʊ	hʊ
JI, JA₁	jɔ	aj		jala	ʊjɛ	uju	ʊnja			ʊjay	jaajʊ	jɪ
KA	kɔ	ak	k	kala	ʊkɛ	uku	ʊnka	k	ka	ʊkay	kaakʊ	kʊ
KU, BUK	(bu)kɔ	ak	k	kala	ʊkɛ	uku	ʊnka	kʊ	ka	ʊkay	kaakʊ	kʊ
MU, MA	mɔ	am	m	mala	ʊmɛ	umu	ʊmma	m	ma	ʊmay	maamʊ	mʊ
NI	ɲɔ	aɲ		ɲala	ʊɲɛ	uɲu	ʊɲɲa		ɲa	ʊɲay	ɲaaɲʊ	ɲɪ
SI	sɔ	as	s	sala	ʊsɛ	usu	ʊnsa	sɪ	sa	ʊsay	saasʊ	sɪ
U	wɔ	aw		wala	ʊwɛ	uwu	ɔwwa	ʊ		ʊway	waawʊ	ʊ
TI (loc)	tɔ	(ad)		tala	ʊtɛ	utu	ʊnta			ʊtay		tɪ
DI (loc)	dɔ	(əd)		dala	ude	udu	undə			udəy		

Figure 2: Non-alliterative class agreement pattern in Keerak

NB: NOM in the table refers to noun classes

16

Considering the issue from a semantic point of view, the analysis of semantic motivation of the distribution of nouns in noun classes in Niger-Congo languages raises the same type of difficulties as the analysis of semantic motivation of the distribution in genders in Indo-European languages. There is no evidence that noun classes in Niger-Congo languages are related to any conceptual classification of meanings in a more direct way than Indo-European genders. "Certain semantic types of nouns tend to cluster in certain classes, but these features are mostly only of statistical value[10]" (Creissels, 1999, p. 181).

Thus, "gender is the intersection of two domains, namely noun classification and syntactic agreement, as the overt expression of a "trigger" (also called controller), usually a noun, on another word as the "target"" (Güldemann & Fiedler, 2019, p.96). But when it comes to going further and defining the main common characteristics of noun classes, there is no overt consensus even though Creissels observes: "recurrent regularities across language families belonging to the Niger-Congo ensemble in which languages with well-developed class systems are found. »
Researchers also agree on the presence of affixes as characteristic of noun classes. Thus, in a noun class system, names are marked by an affix. Although opinions differ as to their functioning. Accordingly, for Gutman and Avanzati: "Nouns are grouped in different classes, marked by prefixes, suffixes or both. All members of a given class share the same affix. Most noun class systems have an accompanying concord system: other elements of the noun phrase (such as determiners, adjectives, or quantifiers), and frequently verbs as well, are marked by an affix selected according to the class of the noun" (Gutman & Avanzati, 2013).
John T. Bendor-Samuel goes into more details by describing how affixes work in Niger Congo languages: "usually one affix signals a singular form and another one signals a plural form. Since these affixes cannot be predicted by phonological or semantic factors, all nouns have to be assigned to classes on the basis of their singular and plural forms. Affixes may be prefixes or suffixes or both, and the number varies from language to language. Most noun class systems have an accompanying concord system, i.e., other elements in the clause -particularly other elements within the noun phrase itself, such as determiners, adjectives, or numerals and frequently verbs- also are marked by an affix selected according to the class of the noun. Similarly, there are sets of pronouns, and the selection of the pronoun is determined by the class to which the pronoun refers. Frequently the same syllable that marks the noun is repeated with other elements; or, if not the identical syllable, a form that has a phonetic resemblance to it is instead repeated" (Bendor-Samuel, 1989).

---

[10] « Certains types sémantiques de noms ont tendance à se regrouper dans certaines classes, mais ces tendances n'ont généralement qu'une valeur statistique » (Creissels, 1999, p. 181)

Finally, semantics brings certain analyses closer when it comes to understanding noun classes and their characteristics. Dixon best defines its contours. For Dixon, it is not enough to study them from a grammatical point of view. On the contrary, one should adopt a "'semantics prior' approach and look to see how semantic types are mapped onto available grammatical choices, then a principled semantic basis may become apparent. "(Dixon, 1986)

According to Gutman and Avanzati (2013), in noun class systems there are several classes, including one semantic group of classes (humans, animals, plants, parts of the body) and others that are based "on grammatical categories but many are heterogeneous" (Gutman and Avanzati, 2013). They are joined by Pozdniakov, who comes to the same conclusion: "genders and classes have different semantical properties (genders distinguish sex whereas classes are rather structured according to animate/non-animate features, as in Niger-Congo languages). Classes would not be complete without semantics: "class assignment is governed by semantic principles so that classes could be described as semantic networks" (Kießling, 2013, p.44).

For Grinevald (1999), it is the manifestation of the affix on the noun and the system of agreement on other constituents that represent the typical characteristics of noun classes.

Nicole (1999) considers that practically every characteristic of noun classes given by Dubois' Dictionary of Linguistics is "false or partially false" and prefers the typological approach of Dixon (1986) that "encompasses all [noun] classification systems". He goes further than Grinevald and specifies that this agreement mechanism can vary from one language to another and that every noun belongs to at least one noun class.

Gutman and Avanzati (2013), for their part, agree on the fact that nouns are grouped into different classes marked by affixes. All members of the same noun class share the same affix. The number of noun classes varies from one language to another, some are semantic and others grammatical. They go further by specifying that: "Most noun class systems have an accompanying concord system that impacts other elements of the noun phrase."

Creissels (2001) agrees with Pozdniakov on the agreement mechanism as a key feature of noun classes. Like many researchers, his analyses are based on the general approach to noun classification. He goes deeper by specifying how they work: "nouns must be divided into a number of subsets to account for the agreement mechanisms between nouns and their modifiers, between nouns and the pronouns that represent them, and finally between nouns and verbs of which there are arguments; words involved in these agreement mechanisms (nouns,

noun modifiers, pronouns and verbs) necessarily carry affixes (class markers) that determine their behavior from the point of view of agreement; all nouns are concerned by classification, which is basically a classification of nouns rather than a classification that would directly take into account the referents of nouns: the assignment of a noun referring to a given type of referent is lexicalized in the sense that speakers are not free to change the class mark of the noun, as it can happen in "classifier" systems, to bring out various semantic features of the same referent" (Creissels, 2001, p.2).[11]

4. *Definition relevant for a Typological Approach of Noun Classes*

For the purposes of this study, we have adopted the definition recently given by Robert Hepburn-Gray (2018), in which he considers a noun class as: "a unique pairing of noun form and agreement class exhibited in the behavior of a noun", as relevant for our study. We retained this definition, amongst others, because it is typological.

Moreover, this definition incorporates the 4 concepts that characterize noun classes according to Güldemann and Fiedler (2019):
   a) Noun Class Marker (NCM)
   b) Gender
   c) Noun class agreement (AGR/NCA)
   d) Deriflection

Güldemann and Fiedler use a methodological and analytical approach to describe these 4 concepts for a "novel analytical approach to gender"(cf. figure 3).

---

[11]« (a) les noms doivent être répartis en un nombre déterminé de sous-ensembles pour rendre compte de mécanismes d'accord entre le nom et ses modifieurs, entre le nom et les pronoms qui le représentent, et enfin entre le nom et le verbe dont il est un argument ; (b)les mots impliquées dans ces mécanismes d'accord (noms, modifieurs de nom, pronoms et verbes) portent obligatoirement des affixes (les marques de classe) qui déterminent leur comportement du point de vue de l'accord (c) tous les noms sont concernés par la classification, qui est fondamentalement une classification des noms plutôt qu'une classification qui prendrait directement en compte les référents des noms : l'affectation d'un nom renvoyant à un type donné de référent est lexicalisée au sens où les locuteurs n'ont pas la liberté de faire varier la marque de classe du nom, comme cela peut se produire dans les systèmes de « classificateurs », pour faire ressortir divers traits sémantiques d'un même référent » (Creissels, 2001, p.2)

Relates to	Concrete noun in a morpho-syntactic context = word form	Abstract noun in the lexicon = lexeme
Syntax	a. AGREEMENT CLASS (abbreviated as AGR and numbered by Arabic numbers)	b. GENDER (numbered by Roman numbers)
Morpho(phono)logy	c. NOMINAL FORM CLASS (abbreviated as NF)	d. DERIFLECTION

Figure 3: The 4 concepts used for analyzing Gender (Güldemann & Fiedler, 2019)

As Hepburn-Gray explains, this definition, by its refined notion of noun classes "which acknowledges both similarities and differences in gender", seems of the greatest interest for the descriptive study that we will carry out in this paper.

Hepburn-Gray study

To better understand Hepburn-Gray's definition of noun classes, we must take into consideration the main lines of his research which have largely inspired the work we will carry out in this paper.

In a survey based on 16 languages (15 modern languages plus Proto-Bantu), representatives of 9 Niger-Congo groups of languages (cf. figure 4), Hepburn-Gray highlights the contribution of his new approach (Hepburn-Gray, 2016). This method, applied to the whole panel of languages, allows him to analyze the main elements, relevant to the notion of noun class, for each language he studied (cf. figures 4 and 5).

Group	Sub-Group	Language	Source
Adamawa	Waja-Jen	Longuda	Jungraithmayr (1968/9)
Atlantic 1	Fula	Gambe Fula	Arnott (1970)
	Nyun	Baïnounk Gubëeher	Cobbinah (2013)
Atlantic 2	Mel	Kisi	Childs (1995)
Benue-Congo	Bantu	Proto-Bantu	Meeussen (1967) & Maho (1999)
	Grassfields	Aghem	Hyman (1981)
	Beboid	Noni	Hyman (1979)
	Kainji	Cicipu	McGill (2007)
Gur	North Central	Ditammari	Reineke (2012)
	Senufo	Supyire	Carlson (1995)
Kordofanian	Heiban	Otoro	Stevenson (2009)
Kru	Eastern	Godié	Marchese (1989)
Kwa	Ka-Togo	Tuwuli	Harley (2005)
	Na-Togo	Logba	Dorvlo (2008)
Ubangi	Mba	a-Mã-lo	Pasch (1986)
	Mba	Ndunga-le	Pasch (1986)

Figure 4: Panel of languages studied by Hepburn-Gray

Language	NCM	T.Aff.	Cls	Gen	MG	Ser.
Otoro	Pre	Mix	14	14	0	3
Gambe Fula	S+M	Mix	25	20	0	11
Baïn. Gub.	Pre	Mix	30	60	0	3
Kisi	Suf	Mix	7	8	5	6
Ditammari	Circ	Mix	17	11	1	8
Supyire	Suf	Mix	8	5	0	7
Longuda	Suf	Mix	9	6	0	5
a-Ma-lo	Suf	Ø	15	9	14	1
Ndunga-le	Suf	Mix	9	9	3	3
Godié	Suf	Suf	6	4	0	2
Tuwuli	Pre	Pre	9	25	0	4
Logba	Pre	Pre	6	13	0	3
Cicipu	Pre	Pre	9	18	0	2
Noni	Pre	Mix	18	22	1	8
Aghem	Pre	Mix	9	11	11	9
Proto-Bantu	Pre	Pre	19	10	0	5

Figure 5: Overview of Languages Analyses

The definition of Güldemann and Fiedler

It seems appropriate to recall the exact meanings used by Güldemann and Fielder (2019):

a) "AGREEMENT CLASS (= AGR): class of concrete noun forms with identical agreement behavior across all relevant targets,

b) GENDER: class of nominal lexemes (or referents) reflecting the agreement-based classification of nouns,

c) NOUN FORM CLASS (= NF): class of concrete noun forms with identical formal, i.e. morpho(phono)logial properties,

d) DERIFLECTION (> DERIvation + inFLECTION): class of nominal (or other) lexical bases established on account of identical morpho(phono)logical paradigmatic variation."

The definitions of the 4 concepts of Güldemann and Fiedler (2019), highlighted by Hepburn-Gray's study, which we use in this paper, have been slightly modified to simplify and make them more affordable:

Noun Class Marker (NCM):
In this paper, the noun class marker is considered as the noun class morpheme specifically on the noun.

The noun class marker is also called "Nominal Form Class" by Corbett (1991) and Güldemann and Fiedler (2019) who defined Nominal Form Classes as "word forms with identical morphological or phonological properties; they represent the counterpart of agreement classes in the realm of morpho(phono)logy. As shown in the important work by Evans (1997) and Evans et al. (1998), nominal form classes (called there "head classes") can have an intricate relationship to agreement classes well beyond serving potentially as their triggers" (Güldemann and Fiedler, 2019, p. 99).

Agreement class (AGR or NCA):
The morphemes added to the concordial words are the Noun Class Agreement (Güldemann & Fiedler, 2019), also called NCA. The NCA encompasses the different forms each concord will take depending on the NCM. The Noun Class Agreement is what is called "Concord" (cf. figure 6).
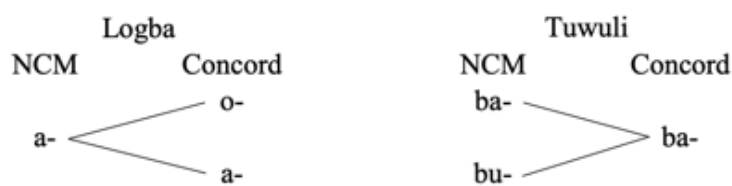
Figure 6: Example of noun classes in Logba and Tuwuli

The NCA is the noun class morpheme on the concordial elements, triggered by the NCM. For example, in Logba, the NCM *a-* can trigger the morpheme *o-* or *a-* on a concord. So, the same NCM can trigger different NCA. It is also possible for two NCMs to trigger the same NCA. For example, in Tuwuli, the NCMs *ba-* and *bu-* trigger the same NCA *ba-*.

Gender:
Hepburn-Gray considers a noun class as a "unique pairing of noun form and agreement class exhibited in the behavior of a noun". In other words, a noun

class is a combination of a noun class form and the noun class agreement it triggers. The gender is the combination of a singular noun class and a plural noun class, both usually sharing semantic features, as it can be seen with the gender 'human' for example, which consists in the pairing of class 1 (SG, human) and class 2 (PL, human). The gender can also comprise only one class if there is no number distinction.

Deriflection:
The term "deriflection", a blend of the words "inflection" and "derivation" was coined by Güldemann and Fiedler (2019), and refers to the interaction between gender and morphology, i.e. it is based on noun classes rather than on agreement and gender.

Deriflection can be witnessed in non-alliterative noun class systems. If a noun class system is alliterative, it means that the NCM will be the same as the NCA. In non-alliterative systems, the NCM and the NCA can differ which shows the existence of two prisms to consider a noun class system : in alliterative only the gender can be used because in this case gender and deriflection are undistinguishable, but in non-alliterative systems a distinction between gender and deriflection needs to be made. The following scheme shows how this distinction can be made.
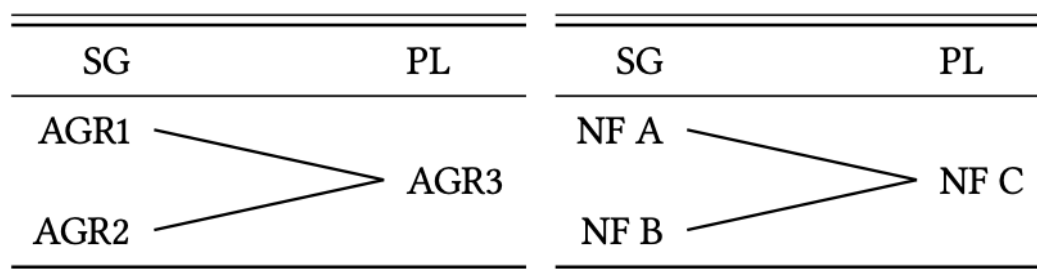
Figure 4: Gender system (left) vs. deriflection system (right) of the case in Figure 3

Figure 7: Gender system and deriflection system – (Güldemann & Fiedler, 2020)

If we follow this scheme, we can explain what is deriflection and when it can be witnessed. In an alliterative system: AGR1 would be triggered by NF A ; AGR2 by NF B ; AGR3 by NF C. The distinction between gender and deriflection is invisible.
In a non-alliterative system : NF A can trigger AGR1 and AGR2 but AGR1 is not necessary triggered by NF A, and it is the same for all the others. In this case, a concrete distinction between gender and deriflection can appear. Two different systems can then be analyzed and the use of links seems necessary to analyze these two systems (cf. Chapter 4, 2. Discussion).

Defining a noun class as "a unique pairing of noun form and agreement class exhibited in the behavior of a noun" as Hepburn-Gray has the advantage of providing a framework that is both broad and precise enough to apply to most of Niger-Congo languages whereas other definitions used to focus on class and gender markers (class matches, unmatched classes, etc.). In fact, noun class systems manifest themselves not only in the morphology of words but also in the syntax of a sentence.

This definition has two benefits: a) in alliterative systems, generalizations can be made with respect to the notion of class b) it removes the need for three distinct characteristics (declension, gender #1 and gender #2) in languages such as Mbembe, and, instead, relies on a combination of two characteristics.

In this approach, noun classes appear as the combination of class markers and the agreement patterns they govern. As Blevins shows, in this conception of noun classes, a predictive dimension is added. It becomes possible to determine plural forms by observing singular forms (Words and Paradigm, Blevins 2013). Noun classes defined as the pairing of noun class form and noun class agreement are "predictive of other paradigms cells" (Blevins, 2013; Hepburn-Gray, 2018). However, it might be necessary to add another parameter to achieve this level of predictability. Indeed, in Bantu languages, classes 1 and 3 are homophonous. For example, in Swahili, class 1 and class 3 share the form *m-*, *mw-*, *mu-*.

This definition of noun classes helps analyzing non-alliterative systems. By combining noun class form and noun class agreement, it becomes possible to predict which noun class should be applied according to grammatical and semantic environment whereas there would be many errors in the prediction of noun class if we only used noun class form.

Conclusion

It appears that noun classes are not the result of an evolution but rather the survival, more or less preserved, of a proto-historical state. Noun classes have been mainly described for Bantu languages because these languages have been the subject of the first and most thorough studies. They were identified as early as 1650 by Giacinto Brusciotto, an Italian missionary, who, at the time, called them "principiatones" (prefixes that determine the rules of agreement). In 1862, Wilhelm Bleek laid the basis for a classification system in which the noun classes of the Bantu languages were numbered. This system initially comprised 16 classes. In 1899, Carl Meinhof presented the first outline of reconstructed proto-Bantu and classified 21

noun classes. In other respects, Sigismund Koelle, comparing 200 languages, attested in 1854 to the existence of noun classes in other African languages subfamilies, typically in the Atlantic languages.

It is difficult to find a simple and comprehensive definition of noun classes in Niger-Congo languages. Researchers use their own definitions or do not define noun classes at all. We can however define some features. A noun is assigned to a class according to its referent, such as gender or form, but this assignment is often purely conventional. Instead of noun classes, some authors use the term "grammatical gender" but others consider that they are different concepts. Nouns are grouped into different classes, marked by prefixes, suffixes or both. All members of a given class that share the same affix. The number of noun classes varies from language to language. Noun classes are usually accompanied by a system of agreement. In the case of an alliterative agreement system, the marker that affects the noun is repeated with other elements such as determiners, adjectives, or numbers. Including in some cases the verb. Nevertheless, in a non-alliterative system, the form of agreement may be different from the affix that affects the noun.

For the purpose of this study, we have adopted the definition given by Robert Hepburn-Gray, in which he considers a noun class as: "a unique pairing of noun form and agreement class exhibited in the behavior of a noun". Based on this definition, we will predominantly use 3 concepts to analyze noun classes in Niger Congo languages: Noun Class Marker (NCM), Agreement class (AGR or NCA), and Gender.

CHAPTER 2: Noun Classes Descriptions

Introduction

The difficulty that linguists face when studying Niger-Congo languages is that noun classes, which are specific to Niger-Congo languages, are not always classified in the same way. They do not always have the same framework and some frameworks may be language specific.

What is highlighted in one grammar may not be used in another one. There is no consensus on how to describe and analyze Niger-Congo noun classes. Despite the current state of knowledge of this phenomenon and the number of Niger-Congo languages studied, the description of noun classes is in fact very close to the definitions, i.e. extremely diverse. However, the first chapter of this paper has highlighted the main notions in noun classes: NCM (Noun Class Marker), NCA (Noun Class Agreement), gender (and deriflection). "What we can expect from a typology is not so much to lock languages into rigid categories, but rather to characterize non-exclusive types of functioning" (Delplanque, 2009).

The second chapter of this paper presents existing noun class frameworks, from the most traditional ones, such as Bantu and Atlantic, to the most uncommon ones, such as Kru and finally Gur.

Creating an Excel table: different classifications, different descriptions

Faced with this variety of classifications, the need appeared for a better understanding of the structures of Niger-Congo noun classes frameworks, and their differences and similarities.

To achieve this goal, the first step was to create a matrix (cf. Table 1), to get an overview of the languages, and their existing noun class classification, and their peculiarities and differences. The matrix details existing classifications for each major Niger-Congo language family. A "convenience sample" of languages, based on availability of sources, has been selected to represent the main families of Niger-Congo languages.

The main purpose of this matrix, built in an excel spreadsheet, presented below, is to highlight some current frameworks for reporting and analyzing noun classes of Niger-Congo languages. It shows whether each language subfamily has a framework for reporting noun classes and thus allows for the aggregation of both commonalities and differences.

In each branch, languages were selected according to the availability of grammars. If more than one grammar was available, the most recent and most complete one was chosen. For some languages, such as Mbembe (Barnwell, 1969) and Supyire (Carlson, 1994), grammars were chosen because of their reputation and, above all, because they have made their mark on African linguistics. Grammars were also selected for their relevance and quality. What characterizes a good grammar is above all its capacity to illustrate the reality of a language rather than to develop a theoretical approach "To create the grammar of a new language [...] is not to create arbitrarily principles according to which that language seems to be spoken, nor to adapt to that language the rules of an already known idiom; but it is to establish theories in connection with the expressions used in practice[12] " (Daull, 1879). Obviously, for some languages, there was no choice, since only one grammar was available. It is therefore necessary to draw the reader's attention to some choices that were made by default, due to the lack of accessibility, but also due to the lack of languages studied. Therefore, the table we have built cannot be considered totally relevant for establishing general rules, but it is a starting point for raising questions and challenges. It aggregates and relates different information about the description of noun class system in each grammar. Each column represents a characteristic necessary for a global vision of Niger-Congo noun classes descriptions.

The 1st column (A) lists the name of the languages being studied and the family (color) to which they belong. There are 9 language families: Kordofanian, Atlantic, Kru, Gur, Adamawa-Ubangi, Kwa, Benue-Congo, Bantoid, Bantu. In most cases, when possible, 4 languages were chosen. The main difficulty, encountered for some families, was the scarcity of languages which have a complete grammar.

The 2nd column (B) mentions the date of publication of the grammar. This input allows to see if the results are influenced by the period in which the grammar was written. A possible impact was highlighted.

---

[12] *« Composer la grammaire d'une langue nouvelle [...], ce n'est pas créer arbitrairement des principes d'après lesquels cette langue nous paraît devoir être parlée, ni non plus adapter absolument à cette langue les règles d'un idiome déjà connu ; mais il s'agit d'établir des théories en rapport avec les expressions reçues dans la pratique »*

The 3rd column (C) indicates the name given to noun classes or the system of noun classes in each grammar. This gives a first idea of what the grammar considers a noun class. For example, noun classes in Wolof (Atlantic), noun genders in Supyire (Gur), and the concord class system in Mbembe (Benue-Congo).

The 4th column (D) lists the labeling system adopted by the relevant grammar. For example, the Bantu tradition uses a numbering system. Some languages will also use number labelling without following the numbering from Bantu Tradition/Proto-Bantu. Other languages, as it is frequent in Atlantic languages, will refer to each noun class using their morpheme. Finally, some grammars display mixed labelling, using numbers and morphemes to refer to different noun classes.

The 5th column (E) is about what the grammar focuses on, more precisely, what seems to be important for the noun class system. The assignment can be semantic, phonological or, rarely, morphological. This information is either clearly indicated or can be deducted from the information made available by the author. In several grammars, the emphasis is on semantics as well as phonology. For example, in Kru languages, grammars focus more on phonology than on semantics. A general indication can be helpful, but in some systems, it may not be possible to state one. Rather, individual noun classes within a language can differ in their main way of categorizing its member nouns (e.g. one NC may be semantically defined, another one formally defined).

The 6th criterion (F) is one of the most important of the table because it relates 'what appears'. When describing noun class systems, linguists usually make a choice in what will be reported and analyzed. This column shows the way noun classes were documented. It can also highlight the reasons why only specific features were discussed but not others when we combine the information from this column with other columns information. What can appear in the description of noun class system is: NCF (NCM), NCA, gender. As described in the first chapter (cf. Chapter 1, 4. Definition relevant for typological approach) of this paper, deriflection will be understood by combining columns *(What appears + Alliterative)*.

The 7th criterion (G) shows whether each grammar contains a table or not. Knowing that the important thing is not so much the layout as the availability of information. A description without a table being more informative than a description with a table that does not contain all the necessary information. This criterion, as it allows to go further on the importance of creating a typological framework for each language, will be important in the third chapter dedicated to challenges.

The 8th criterion (H) lists whether there is an agreement pattern or not, and if this system is alliterative or not.

	Date Grammar	Refers to noun class system as	Labelling system	Focus on	What appears	Table ?	Alliterative agreement system	Commentaires
KORDOFANIAN								
Lumun	2017	Noun classes	morpheme	Semantics	NCF (NCM) ; NCA	Yes	yes	"subclasses" ; consonantic morpheme ; The focus on semantics is not that strong
Dagik	2016	Noun class system	morpheme	Semantics	Gender ; NCA	No (Links)	yes	consonantic morpheme
Heiban	1997	Noun classes	morpheme	Semantics	NCF (NCM) ; NCA	Yes	Non alliterative	Most of the concord agreements are alliterative but not all of them.
Moro	2009	Noun class system	morpheme	Phonology ; Semantics	NCF (NCM) ; NCA	Yes (+links)	yes	
ATLANTIC								
Kisi	2011	The noun class system	morpheme	Semantics	NCF (NCM)	No	Non alliterative	Adjectival concord ; only shows the NCA in examples and only for adjectival concord.
Yoruba								No noun class system anymore
Wolof	2015	Les classes nominales	Morpheme	Phonology ; Semantics	NCF (NCM) ; NCA	Yes	Non alliterative?	
Wolof	2016	Noun classes	Number ; morpheme		Gender ; NCM ; NCA	No (Links)		NCM inexistant --> NCM fused with definite or indefinite. The traditional subdivision of Wolof nouns into different NCs relies exclusively on determiner agreement as shown in (1): in fact, selection of one form of the determiner, picked from a set of possible alternatives within its paradigm (cf. (13) below), fits the definition of agreement as "systematic covariance between a semantic or formal property of one element and a formal property of another" (Steele 1978: 610). In other Niger-Congo (including Atlantic) languages, on the other hand, not only is there NC-agreement, but also nouns themselves are class-marked, as exemplified by Diola-Fogny in (3) (Sapir 1965: 24, 90)
Keerak	2015	Les classes nominales	morpheme	Semantics	NCA	Yes	Non alliterative	"A ce stade de la description, les classes sont désignées par des majuscules et sont définies sur la base des préfixes nominaux, indépendamment de leur schème d'accord, de leur valeur pour ce qui concerne le nombre et de leur contenu sémantique"
Temne	2007	Noun classes	Number ; morpheme	Semantics	NCF (NCM)	Yes	Non alliterative	
Noon	2000	The noun class system	Number	Semantics	NCM ; NCA	Yes		The class corresponds to the Gender : each class as a different form depending on the number (sg/pl) ; No use of the word gender ; Pour class 1 : "wherever a noun and a verb share a root, the noun belongs to this class".
KRU								
Dadjriwale	2008	Le pluriel des classes non-humains	"humains" / "Non-humains"	Phonology ; Semantics	NCA	No	Non alliterative	Dans ce dialecte : 2 classes nominales (humains / non-humains) ; utilise le terme de classes nominales mais jamais comme titre. Pour NCA, une section pour chaque concord montrant les différents accords. Aucun tableau réacipitulatif.
Bassa	2012							No Noun class ?????? Gender distinction : masculine/feminine
Bete	2004	Le pluriel	Voyelle	Phonology	NCM	No		"Parlons Bété" n'est pas une description/grammar ==> teaches you how to speak Bete ; never uses the term noun class ; "Les noms propres réalisent leur pluriel en utilisant le nominatif /ngwa/ suffixé au nom" "Pour les autres mots, noms et sous-classes de nom, la formation du pluriel est de type vocalique. La voyelle finale du mot au pluriel dépend de la voyelle finale du mot au singulier."
GUR								
Koulango	2016	Suffixes de classes nominales	morpheme	Phonology ; Semantics	NCF (NCM) ; NCA	No	Non alliterative	"schèmes d'accord" pour le défini, le qualificatif et le pronom, les différents genre sont animé, inanimé, mais c'est pas vraiment dit.
Supyire	1994	Noun genders	Number (gender)	Phonology ; Semantics	Gender ; NCM ; NCA	Yes	Non alliterative	
Tagbana	2018	Nominal classes	Number		(comments*)		Gender ; NCM ; NCA	Number for classes and for genders. (4 genders, 7 classes) ; *Focus ni sur la sémantique ou sur la phonology mais sur l'agreement ... Syntaxe ?
Kusaal	2018	Noun class system	Number	Semantics	NCM	Yes (+links)		Pairing are established but never called gender. The word gender does not appear. There is no way to tell if it is an alliterative system or not.
Chakali	2017	Noun classes	Number	Phonology ; Semantics	NCM	Yes		"class can be regarded as phonological and/or semantic features encoded in the lexemes for the selection of the proper pair of singular and plural suffixes." Le NCA n'est pas discuté du tout on sait pas
Biali	2014	Système des classes nominales	Roman numerals	Semantics	Gender ; NCM ; NCA	Yes	Non alliterative	Roman numerals for gender, nothing for the classes
(Gidere)								
ADAMAWA-UBANGI								
Mbodomo	1997							No Noun class ??????
Mundang		Vestiges de classess nominales						
Samba Leko								
(Mambay)								"Noun prefixes in Mambay do not appear to be related to the systems of noun classification prevalent elsewhere in Niger-Congo (5.1.2, Anonby 2008:3; contra Greenberg 1963:9, 11)"
KWA								
Tuwuli	2005	Noun classes	Number (Class/gender, letters (subclasses)	Phonology ; Semantics	NCF (NCM) ; NCA	Yes	Non alliterative	"Noun classes are determined by the plural prefixes alone, at least in those lasses which have an inflectional singular/plural distinction." Classes are considered as gender ; subclasses. NCA sometimes mandatory, sometimes not ? Tableaux p.104 et 107
Baoule	1977							By Creissels : pas des classes nominales mais classes nominales?????
Efutu	2008							There seems to be prefixal noun class but the writer does not analyze them that way.
Tafi	2013	Noun classes	Morpheme	Phonology ; Semantics	NCF (NCM)	Yes		Numbers in the table but never used in the paper. Pairing (sg/pl) with links, but never used the term gender. "do not show agreement between nouns and their qualifiers."
Logba	2008	Noun classes	morpheme	Phonology	NCF(NCM) ; NCA	No	Non alliterative	Ne montre pas les liens entre NCM et NCA. "The nouns which have the o-/ɔ- verbal concord are by far the largest group of nouns. They comprise nouns with the following prefixes. o-/ɔ-, u-, e-/e-, and a- prefix nouns18. These are all singular nouns. Nouns that trigger the i- verbal concord are those that take i- noun prefix. Nouns whose stems belong to the o-/ɔ- noun prefix take the o-/ɔ- sin- gular class. Nouns that are cross referenced by the N- prefix as verbal concord are those nouns that take the N-noun prefix. Nouns with the e-/ɛ- plural noun prefix trigger the concord of the same form." Focuses mostly on phonology, very rarely talks about the semantics. However it seems to be related : p. 70-71 (48-49)
(Gidere)								
BENUE-CONGO								
Mbembe	1969	The concord class system	number ("sets")		NCF (NCM) ; NCA		Non alliterative	
BANTOID								
Nchane								
Mundabli	2017	Noun class system	Number	Semantics	Gender ; NCF (NCM)	Yes (+links)		
Mungbam	2013		Number			Yes (+links)		
BANTU								
Swahili								
Tswana								
Liko	2015	Noun classes	Number	Phonology ; Semantics	NCF (NCM) ; NCA	Yes	Non alliterative	
Kikongo								
Lingala	2010	Noun class system	Number	Semantics	NCF (NCM) ; NCA	Yes	yes	The author also uses the morpheme (not only the number), bu the numbers have the value given by Bantu tradition. He talks about the pairing but never qualify this pairing as a gender. Due to phonological reasons, the prefix can have different forms.
(Bila)								
Isangu	1998	Les préfixes nominaux et classes d'accord	Number	Phonology ; Semantics	Gender ; NCM ; NCA	Yes	Non alliterative	"Le préfixe marque le nombre et le genre du nom". Il fait référence à Guthrie : *Comparative Bantu* , Volume II ; il y montre les classes données par Guthrie : 1. mu-, 2.ba-, etc. L'auteur utilise tout de même le terme de classes.
Bemba	2009	Les classes nominales	Number	Phonology ; Semantics	Gender ; NCM ; NCA	Yes	Non alliterative	Compares the NCM with Proto-Bantu. Phonology. Surtout syntxe pour déterminer quelles sont les classes. La sémantique n'est évoquée uniquement pour les genres. NCM, Gender et NCA sont traités indépendamment et n'apparaissent jamais les trois simultanément dans un tableau.
Makwe	2008	Noun classes	Number	Phonology ; Semantics	Gender ; NCM ; NCA	Yes	Non alliterative	Also uses the morphemes. Phonology : Certains morphèmes vont prendre une forme différentes pour des raisons phonologiques (différentes raisons). Ils font toujours partie de la même classe nominale (je crois). Links between noun classes are made (genders and agreement patterns), but not in a table.

Table 1: Niger-Congo noun classes descriptions

In this second chapter, we investigate the main branches of Niger-Congo languages, using our Excel table. Each section of this chapter is designed according to the same pattern: after a review of the noun class system of each branch and how it was previously studied, we focus on what is apparent in our Excel table and on what is relevant in this classification. Then, we set the typological limitations of each classification. Finally, when possible, we attempt to highlight similarities between different subfamilies frameworks of noun class systems.

1. *Bantu Descriptions*

This review of Niger-Congo languages will begin with Bantu languages. The Bantu tradition is perhaps the most influential, as it is the largest family in terms of the number of languages spoken (between 400 and 500) and in terms of geographical extent (half of Africa). This makes it one of the most studied language families. For Creissels (2001), "the most typical languages" from the point of view of noun classes "are Bantu and Atlantic languages". Moreover, "the notion of noun class is the major characteristic of Bantu languages, indeed of the entire Niger-Congo phylum, since more than two thirds of the languages of this family are concerned by this feature"[13] (Mouguiama-Daouda, 2005). Moreover, Creissels (2015) considers that "by their general architecture, Bantu noun class systems are on the whole typical of Niger-Congo noun class systems". It was therefore essential to devote significant space to them. This is why our Excel table includes 9 Bantu languages in order to give a global view of this branch (Swahili, Tswana, Liko, Kikongo, Lingala, (Bila), Isangu, Bemba, Makwe). In particular, Swahili and Lingala are representative as they have quite different noun class systems. Not in terms of prefixes on the noun, but in terms of agreement, and what triggers it.

The first chapter of this paper showed that Bantu noun classes were among the first to be discovered and analyzed (cf. Chapter 1, 1. Historical context). The first frameworks (cf. Meinhof-Bleek / Bleek Meinhof) served as the basis for the current frameworks used for Bantu noun classes. "The classification of African languages in terms of "noun classes" was initiated by the German William H. J. Bleek (1827-1875) and adopted by the German Carl Meinhof (1857-1944)" who, in turn, highlighted the 21 noun classes of Bantu languages (Kantchoa, 2014). We can assume that few changes have been made since then. The same framework is still used, linguists have only added new types of noun classes, such as kinship declension, numbered class 1a or 2a in some cases. Lumwamu (1970) believed that Van Bulk's Handbook of Bantu Linguistics (1948) lists them as a fundamental work that is commonly used.

1.1.   *Framework description*

The traditional noun classes Bantu framework is based on the proto-Bantu noun class system: Meinhof (1932), Meeussen (1967), Welmers (1973), Hinnebusch (1989) (Maho, 1999, p.51).

---

[13] « La notion de classe nominale est la caractéristique majeure des langues bantu, voire de toute la branche niger-congo, puisque plus de deux-tiers des langues de cette famille sont concernés par ce trait. »

This approach consists of assigning a class number to nouns, generally ranging from 1 to 20, and, sometimes, adding letters – the corresponding morpheme and the semantic of the relevant class. Still, they are not decisive for the analysis on the same level. In some descriptions, for instance, semantics is added as additional info about semantic tendencies, but it is not used as a crucial factor in deciding what noun class a noun belongs to (or very rarely does so).

Each number corresponds to a specific meaning when letters refer to a function. Singular and plural are considered as separate semantic features. Based on the example of Tswana (cf. Figure 8), the illustration below shows the characterization of Bantu, thus Niger-Congo, noun class systems. We note a division of: " noun forms into classes in which the singular form and the plural form of the same noun are treated as two distinct units, the notion of gender being subsequently introduced as a pair of classes that correspond to each other in the expression of number"[14] (Creissels, 2001, p.2).

a.	cl. 1	**mo**-sadi **yo mo**-ša	'nouvelle femme'
b.	cl. 2	**ba**-sadi **ba ba**-ša	'nouvelles femmes'
c.	cl. 3	**mo**-lemo **o mo**-ša	'nouveau médicament'
d.	cl. 4	**me**-lemo **e me**-ša	'nouveaux médicaments'
e.	cl. 5	**le**-saka **le le**-ša	'nouvel enclos'
f.	cl. 6	**ma**-raka **a ma**-ša	'nouveaux enclos'
g.	cl. 7	**se**-kolo **se se**-ša	'nouvelle école'
h.	cl. 8-10	**di**-kolo **tse din**-tšha	'nouvelles écoles'
		di-kgosi **tse din**-tšha	'nouveaux chefs'
i.	cl. 9	**ø**-kgosi **e n**-tšha	'nouveau chef'
j.	cl. 11	**lo**-kwalo **lo lo**-ša	'nouveau livre'
k.	cl. 14	**bo**-jang **jo bo**-ša	'nouvelle herbe'
l.	cl. 15-17	**go** lema **mo go**-ša	'nouvelle façon de cultiver'

Figure 8: Tswana noun class system – (Creissels, 2001)

Moreover, Creissels (2001, p. 1) notes that "noun class systems of Niger-Congo languages are both typologically close to 'gender' systems and clearly different from 'classifier' systems: like 'gender', they involve highly grammaticalized agreement mechanisms that have no equivalent in 'classifier' systems"[15]. This is

---

[14] "*des formes nominales en classes dans laquelle la forme de singulier et la forme de pluriel d'un même nom sont traitées comme deux unités distinctes, la notion de genre étant ultérieurement introduite comme couple de classes qui se correspondent dans l'expression du nombre*" »

[15] "les systèmes de classes nominales des langues Niger- Congo sont typologiquement à la fois proches des systèmes de « genre » et nettement différents des systèmes de « classificateurs » :

considered highly grammaticalized because an agreement in noun class is mandatory for the dependent modifiers. For example, in Tswana, we can see that each noun modifier receives a noun class marker to show the agreement (Creissels, 2001):

a. mo-sadi yo mo-leele yo mo-ntsho
1-femme jonctcl1 1-grand jonctcl1 1-noir
yo o opelang yo-le
jonctcl1 scl1 chanter.rel 1-dém
'cette femme de haute taille et au teint foncé qui chante'
b. le-kau le le-leele le le-ntsho
5-garçon jonctcl5 5-grand jonctcl5 5-noir
le   le opelang le-le
jonctcl5 scl5 chanter.rel 5-dém
'ce garçon de haute taille et au teint foncé qui chante'

Figure 9: Tswana noun class system – (Creissels, 2001)

1.2.   The contribution of the Excel table

BANTU								
Swahili								
Tswana								
Liko	2015	Noun classes	Number	Phonology ; Semantics	NCF (NCM) ; NCA	Yes	Non alliterative	
Kikongo								
Lingala (Bila)	2010	Noun class system	Number	Semantics	NCF (NCM) ; NCA	Yes	yes	The author also uses the morpheme (not only the number), bu the numbers have the value given by Bantu tradition. He talks about the pairing but never qualify this pairing as a gender. Due to phonological reasons, the prefix can have different forms.
Isangu	1998	Les préfixes nominaux et classes d'accord	Number	Phonology ; Semantics	Gender ; NCM ; NCA	Yes	Non alliterative	"Le préfixe marque le nombre et le genre du nom". Il fait référence à Guthrie : *Comparative Bantu* , Volume II ; il y montre les classes données par Guthrie : 1. mu-, 2.ba-, etc. L'auteur utilise tout de même le terme de classes.
Bemba	2009	Les classes nominales	Number	Phonology ; Semantics	Gender ; NCM ; NCA	Yes	Non alliterative	Compares the NCM with Proto-Bantu. Phonology. Surtout syntxe pour déterminer quelles sont les classes. La sémantique n'est évoquée uniquement pour les genres. NCM, Gender et NCA sont traîtés indépendamment et n'apparaissent jamais les trois simultanément dans un tableau.
Makwe	2008	Noun classes	Number	Phonology ; Semantics	Gender ; NCM ; NCA	Yes	Non alliterative	Also uses the morphemes. Phonology : Certains morphèmes vont prendre une forme différentes pour des raisons phonologiques (différentes raisons). Ils font toujours partie de la même classe nominale (je crois). Links between noun classes are made (genders and agreement patterns), but not in a table.

Table 1.1: Bantu noun classes descriptions

Lingala and Swahili are representative of the Bantu framework in our excel table. By comparing the different ways of studying Bantu languages, it shows that there is a general consistency and homogeneity in the way linguists report on Bantu noun classes, and that they tend to focus on the same features. Moreover, studies typically include a table and focus on NCF (NCM) and NCA. We also note that gender is not highlighted, and that it does not appear as such, though it is well known that noun

---

comme le « genre », ils mettent en jeu de manière cruciale des mécanismes d'accord fortement grammaticalisés qui n'ont pas leur équivalent dans les systèmes de « classificateurs ». " (Creissels, 2001, p.1)

class pairings will be considered as gender. Gender is acknowledged in the Bantu tradition, but it is not overtly shown, and it is rarely shown in tables.

On the other hand, a strong focus on NCM and NCA is made which is very helpful for languages with non-alliterative noun classes morphemes on each agreeing category. In the case of the Bantu languages selected in our Excel table, focusing on NCA makes sense. Whether or not agreement is alliterative differs from NC to NC, as well as from target category to target category in some languages (e.g. Swahili *ki-* (NC7) is alliterative, vs. *n-* (NC9): allit. on ADJ, but non-allit. on verbs). The focus on NCM and NCA is very useful for languages with non-alliterative noun class morphemes on each agreement.

Semantics are present all the time and always considered regardless of the language and of the linguist. Indeed, noun classes of proto-Bantu languages (and their number) are separated on a semantic basis, they use semantic to be differentiated (other features can be used too).

This Excel table also highlights similarities between Bantu, Bantoid and Benue-Congo classifications (ways of reporting noun classes). These two latter classifications also label noun classes with numbers. However, these numbers do not always correspond to the Proto-Bantu numbers. A correlation between these three frameworks is not such a surprise given their genealogical link.

1.3. (Typological) limits of the Bantu's framework

As discussed previously, gender is not a proper category in the Bantu's framework but is often acknowledged. The way to signify gender is through the pairing of noun classes which can also lead to misunderstandings. For instance, class 1 and 2 refer to the same meaning but have a different number and class 1 and 7 refer to different referents but share the same number. In other words, the fact that gender does not overtly appear makes a typological analysis of gender impossible, or very fastidious, with a high proportion of mistakes. As explained later in this paper, the non-overt gender also entails a lack of distinction between inflectional and derivational morphology in noun class systems. Thus, number is considered as a semantic feature and not as a grammatical feature.

Generally speaking, that the tables analyzing Bantu languages are very rigid because they are specific to Bantu languages and therefore not flexible enough to be adapted to other Niger-Congo languages. We will see that this rigidity and therefore this problem is encountered with all the tables used to study other Niger-Congo languages.

1.4. Typological advantages

The Bantu tradition is a common knowledge among Africanists which enables a concrete consistency in noun classes. By following the Proto-Bantu numbers, each noun class system will be illustrated in a common manner which makes this approach comparable across languages. Each class possesses a specific value. In some languages, some of these values are not used to categorize the world around the speakers (gender), therefore the class referring to the absent value in this language will not appear in the noun class system table. For example, in Swahili, not all the noun classes from Proto-Bantu appear, classes 12, 13, 14, 19, 20, 21, (22,) and 23 do not appear[16]:

Classe	Noms	Adjectifs	Démons.	Possess.	Sujet	Objet
1	m(w)-	m(w)-	yu-	w-	a-, (yu-)	-m(w)-
2	wa-	wa-	wa-	w-	wa-	-wa-
3	m(w)-	m(w)-	u-	w-	u-	-u-
4	mi-	mi-	i-	y-	i-	-i-
5	Ø-, ji-	Ø-, ji-	li-	l-	li-	-li-
6	ma-	ma-	ya-	y-	ya-	-ya-
7	ki-	ki-	ki-	ch-	ki-	-ki-
8	vi	vi-	vi-	vy-	vi-	-vi-
9	N-, Ø-	N-, Ø-	i-	y-	i-	-i-
10	N-, Ø-	N-, Ø-	zi-	z-	zi-	-zi-
11	u-	m(w)-	u-	w-	u-	-u-
15	ku-	ku-	ku-	kw-	ku-	-ku-
16	--	pa-	pa-	p-	pa-	-pa-
17	--	ku-	ku-	kw-	ku-	-ku-
18	--	m(w)-	m(u)-	mw-	m(w)-	?

Figure 10: Swahili noun class system – (Philippson, 2017)

The consistency and correspondence in the classes help to consider this framework as typological though it is not typological. It offers a certain cross-linguistic consistency that can sometimes allow for typological research though the basis of the Bantu framework is not typological but genetic, mainly based on diachronic forms.

2. Atlantic Descriptions

As mentioned above, we are in the process of reviewing the different noun class frameworks in the major Niger-Congo language families. This section focuses on the Atlantic family which is composed of more than 60 languages spoken by more than 30 million speakers in a dozen countries such as Burkina Faso, Cameroon, Gambia, Guinea, Guinea-Bissau, Liberia, Mali, Mauritania, Niger, Nigeria, Senegal, Sierra Leone or Chad. Pozdniakov and Ségerer (2017) separate the Atlantic languages into two groups, north (N) and central (C), as they consider them to have

---

[16] Gérard Philippson, Cours de swahili : Le système de classes nominales - Komedit (septembre 2017)

different linguistic behaviors. Five Atlantic languages are analyzed in the Table 1.2: Kisi (C), Wolof (N), Keerak (N), Temne (C) and Noon (N).

Creissels (2015) notes that "the noun class systems of Central Atlantic languages remain closer to the Niger-Congo prototype than those of Northern Atlantic languages. "For example, "Noun class systems in which nouns (or a portion of them) lack a noun class marker are found only in two groups of languages in the North Atlantic branch (Wolof and Cangin)." Unlike the Bantu languages (which follow the Proto-Bantu framework), the numbering is arbitrary: "As opposed to Bantu noun class systems, they do not relate to a Proto-Atlantic noun class system" (Creissels, 2015).

2.1.   The contributions of the Excel table for Atlantic languages

ATLANTIC								
Kisi	2011	The noun class system	morpheme	Semantics	NCF (NCM)	No	Non alliterative	Adjectival concord ; only shows the NCA in examples and only for adjectival concord.
Yoruba								No noun class system anymore
Wolof	2015	Les classes nominales	Morpheme	Phonology ; Semantics	NCF (NCM) ; NCA	Yes	Non alliterative?	
Wolof	2016	Noun classes	Number ; morpheme		Gender ; NCM ; NCA	No (Links)		NCM inexistant --> NCM fused with definite or indefinite. The traditional subdivision of Wolof nouns into different NCs relies exclusively on determiner agreement as shown in (1): in fact, selection of one form of the determiner, picked from a set of possible alternatives within its paradigm (cf. (13) below), fits the definition of agreement as "systematic covariance between a semantic or formal property of one element and a formal property of another" (Steele 1978: 610). In other Niger-Congo (including Atlantic) languages, on the other hand, not only is there NC-agreement, but also nouns themselves are class-marked, as exemplified by Diola-Fogny in (3) (Sapir 1965: 24, 90)
Keerak	2015	Les classes nominales	morpheme	Semantics	NCA	Yes	Non alliterative	"A ce stade de la description, les classes sont désignées par des majuscules et sont définies sur la base des préfixes nominaux, indépendamment de leur schème d'accord, de leur valeur pour ce qui concerne le nombre et de leur contenu sémantique"
Temne	2007	Noun classes	Number ; morpheme	Semantics	NCF (NCM)	Yes	Non alliterative	
Noon	2000	The noun class system	Number	Semantics	NCM ; NCA	Yes	Non alliterative	The class corresponds to the Gender : each class as a different form depending on the number (sg/pl) ; No use of the word gender ; Pour class 1 : "wherever a noun and a verb share a root, the noun belongs to this class".

Table 1.2: Atlantic noun classes descriptions

Our Excel table shows that, for Atlantic languages, noun class systems refer to noun classes using morphemes. Indeed, we can see that only the Noon description uses numbers, while other descriptions use morphemes. In Noon, numbers are arbitrary, they do not follow any typology or genealogy or tradition. This is also the case for the other languages that use numbers: the first classes refer to humans, but most of the others have no correlation with each other. Morphemes are also used because linguists did not want to arbitrarily establish apparent priority. By using morphemes, it was clear that it was very specific to the language. It is important to add that numbers in Atlantic noun classes descriptions would not correspond to the numbers used in the Bantu tradition.

A table summarizing the noun class system of each language is not systematically present in grammars. The representation of noun classes is always different. Even if tables are used, 'what will appear' will not always be the same. For most of them, NCM will appear. However, Keerak will only show the NCA. Gender (concept) is only represented in Wolof, which has been studied more than other Atlantic languages.

The emphasis on NCA seems to be quite general for Atlantic languages, which can be explained by the fact that most of languages have non-alliterative noun class

systems. However, in the descriptions of Kisi and Temne, only the NCM appears. Our table shows that there is no Atlantic common framework defined by any linguist and that there is no way to account for an Atlantic system of noun classes. Unlike Bantu languages where there is an established tradition, for Atlantic languages, we do not detect any framework to follow. Grammars are different but consistent in that they use mainly morphemes, and they focus mainly on NCA.

A compelling example for the absence of an Atlantic noun class framework can be found in Wolof descriptions. Indeed, in the Excel table, two different grammars of Wolof were analyzed. One uses numbers and morphemes while the other one uses morphemes only. One shows the gender when the one only shows the NCM and the NCA.

2.2. Why this framework for this subfamily?

In a very different approach from the Bantu tradition, which appears to be quite coherent, because of the framework provided by the proto-Bantu, in the case of the Atlantic subfamily, there is no real proto-Atlantic reconstruction and therefore no coherence in the framework used by linguists. However, they seem to be aware of the arbitrariness of their framework and most of them decided to go deeper into the specificity, and to document the language. So, they created frameworks that would meet the needs of each language to represent their noun class systems.

2.3. The typological limits of Atlantic frameworks

For this framework to be typological, a first limit seems quite obvious: to refer to noun classes, most linguists have decided to use morphemes, which shows that this framework is not typological but deeply language specific.

There is no generic Atlantic framework for representing noun classes, but if we were to consider that the Atlantic framework is to use morphemes and focus on NCA, then this framework would be too language-specific to be applied to other noun class systems.

2.4. Other language subfamilies with similarities: KORDOFANIAN (according to Table 1)

KORDOFANIAN								
Lumun	2017	Noun classes	morpheme	Semantics	NCF (NCM) ; NCA	Yes	yes	"subclasses" ; consonantic morpheme ; The focus on semantics is not that strong
Dagik	2016	Noun class system	morpheme	Semantics	Gender ; NCA	No (Links)	yes	consonantic morpheme
Heiban	1997	Noun classes	morpheme	Semantics	NCF (NCM) ; NCA	Yes	Non alliterative	Most of the concord agreements are alliterative but not all of them.
Moro	2009	Noun class system	morpheme	Phonology ; Semantics	NCF (NCM) ; NCA	Yes (+links)	yes	

Table 1.3: Kordofanian noun classes descriptions

It is important to remind that Kordofanian languages and Atlantic languages are very different and spoken in very different places in Africa. However, Kordofanian noun classes descriptions have similarities with the Atlantic ones. Indeed, they share the same labeling system based on morphemes. They also share the fact that there is no strong focus on gender as a term, which is only found for one language: Dagik. On the other hand, constant appearance of NCA in each grammar is a bit different with Atlantic noun classes representation. Besides, most of noun classes are alliterative or lightly alliterative.

There is not much literature on Kordofanian noun class systems, which are definitely studied in a quite language specific way. Most of information that were accessible are from grammars of Kordofanian languages, then generalized to the subfamily. "Most Kordofanian languages have noun class systems—that is, systems in which each noun is marked by one of a set of affixes—and other elements in a clause (numerals and adjectives, for example) are also marked by an affix determined by the respective noun class" (Bendor-Samuel, 2012).

3. *Kru Descriptions*

Kru languages are a branch of the Niger-Congo phylum. They are spoken by about 3 million people living in Southwestern Côte d'Ivoire and Southern Liberia. The Kru sub-family counts 24 languages. All of them sharing the tonal feature, which can sometimes be used to distinguish between definite and indefinite or singular and plural. Tones can be used as a distinctive feature for noun classes.
Though there are evidence for a proto noun class system putting animacy as a very important feature, some Kru languages added a feminine category to the noun class system which seems very unusual. "Nonhuman classes in three languages have been re analyzed as feminine, adding a new dimension to an otherwise genderless system" (Marchese, 1988).

"Kru languages are attached, within the Niger-Congo phylum, to the northern Volta-Congo group, which also includes Gur languages and Adamawa-Ubangian languages (Williamson & Blench 2000). Volta-Congo languages are nominally suffixed languages. They differ from Kwa languages (Bolé-Richard 1983) and Bantu languages (Rébuschi 1999) whose affixation mode is prefixation" (N'dré, 2016).

There is little literature available on Kru languages. Although there are a few pieces of description dating back to 1845, it seems that descriptions of these languages are only beginning to be made. There are not many grammars available and the

only grammars that can be found were published recently (between 2004 and 2012).

Taking this into account, our Table 1 only gathers information for two Kru languages. The results must be considered with care and might not be representative of Kru noun class framework. Marchese (1988) encountered a similar problem, "data in this paper come entirely from synchronic sources, there being very little historical documentation on this group".

Marchese, in a Chapter of *Agreement in natural languages* (Barlow and Ferguson, 1988)*,* attempted to reconstruct the Proto-Kru noun class system. He counted four singular and three plural classes and highlighted the noun classes reduction currently taking place in Kru languages.

3.1. Description, literature

Kru noun classes differ greatly from the rest of Niger-Congo noun class systems for two reasons. First, the singular form will always be a morpheme zero, which made it more complicated for linguists to see/detect the gender (the opposition between the different singular forms is not overt). Second, the plural noun classes, which are the only ones overtly appearing, are phonologically conditioned, their form will depend on the ending vowel of the stem. As Godé (2008, p. 101-102) explains in other words: "In Kru languages, there are no formal marks specialized in the representation of singular. This number is marked by a morpheme zero. The plural number has a form, but it is linked to the final vowel of the singular nominal root. This process is more a matter of phonological than grammatical agreement, nevertheless it makes it possible to justify the existence of noun classes in these languages."[17]

In his book giving a typological overview on Kru languages, Rickard (1970) never uses the term "noun classes". He only talks about affixation and lists plural forms of nouns showing that they depend on the ending vowel on the singular word.

According to François Ble Kipré (2016, p.15): "Noun class systems of Kru languages are so deteriorated that they are now reduced to the state of 'noun class remnants' (Marchese, 1983): class markers are not represented by affixes on nouns, as in Bantu

---

[17] "Dans les langues kru, il n'existe pas de marques formelles spécialisées dans la représentation du nombre singulier. ce nombre est marqué par un morphème zéro. Le nombre pluriel, lui, a une forme, mais celle-ci est liée à la voyelle finale du radical nominal singulier. Ce procédé qui relève plutôt d'un accord phonologique que grammatical permet néanmoins de justifier de l'existence des classes nominales dans ces langues. » (Godé, 2008)

languages considered to be prototypes of noun class languages at the level of Niger-Congo. Noun classes are most evident in the language pronominal system. The only semantic distinction obviously related to the formal noun class system is + or - human."[18]

Marchese Zogbo (2005), on her side, analyzed this phenomenon as a "kind of phonological agreement system", but considers this system as a result of the proto class system.
The literature does not seem to offer a clearly defined framework for Kru noun classes. Rickard (1970) decided to explain point by point the anticipated plural form depending on the ending vowel of a noun. As we will see in the Table 1.5 analysis, this listing type of framework, mixed with paragraphs appears to be common in Kru noun classes descriptions.

3.2. Excel Table

KRU								
Dadjriwale	2008	Le pluriel des classes non-humains	"humains" / "Non-humains"	Phonology ; Semantics	NCA	No	Non alliterative	Dans ce dialecte : 2 classes nominales (humains / non-humains) ; utilise le term de classes nominales mais jamais comme titre. Pour NCA, une section pour chaque concord montrant les différents accords. Aucun tableau réacpitulatif.
Bassa	2012							No Noun class ?????? Gender distinction : masculine/feminine
Bete	2004	Le pluriel	Voyelle	Phonology	NCM	No	ngwa	"Parlons Bété" is not a description/grammar ==> teaches you how to speak Bete ; never uses the term noun classes ; "Les noms propres réalisent leur pluriel en utilisant le nominatif /ngwa/ suffixé au nom" "Pour les autres mots, noms et sous-classes de nom, la formation du pluriel est de type vocalique. La voyelle finale du mot au pluriel dépend de la voyelle finale du mot au singulier."

Table 1.4: Kru noun classes descriptions

In the excel table, only two Kru languages are analyzed: Bété, spoken in the regions of Gagnoa and Daloa of Côte d'Ivoire, and Dadjriwalé, one of the dialects of Godié, spoken in the west of Côte d'Ivoire, particularly in the sub-prefecture of Fresco. The interpretation of this table for Kru noun classes descriptions cannot be generalized due to a non-significative amount of data. Bassa, spoken in Central Africa, Cameroon, in the regions of the Centre, Littoral and South, does not use noun classes as a gender distinction but masculine/feminine. This is why is it not analysed.

It seems that there is no defined framework for Kru noun classes, mostly because Kru noun class system is more of a phonological phenomenon than a grammatical one, although it remains a grammaticalized system.

---

[18] "Les systèmes de classification nominale des langues kru sont si détériorés qu'ils sont aujourd'hui réduits à l'état de « vestiges de classes nominales » (Marchese, 1983 :189) : les marqueurs de classes ne sont pas représentés par des affixes sur les noms, comme c'est le cas dans les langues bantoues considérés comme les prototypes de langues à classes nominales au niveau du Niger-Congo. Les classes nominales se manifestent surtout dans le système pronominal de la langue. La seule distinction sémantique reliée de manière évidente au système formel de classes nominale est + ou - humain. "

As we can see from the table, linguists never refer to noun classes as such. Instead, they only refer to plural or plural of non-human classes. However, the Dadjriwalé description (Godé, 2008) acknowledges the existence of noun classes, but does not use this term in the title.

Phonology is central to the assignment of noun classes in Bété. However, semantics can sometimes be taken into account in Dadjriwalé. For example, in Dadjriwalé, there are only two classes, and Godé has opted for labels such as 'human' and 'non-human'. The labeling system is language specific, either because there are only a few noun classes, or because noun classes are not labeled. For example, in Bété, only morphemes are used, but not really as labels. Because noun classes are not considered as noun classes in this Bete description, there is no real labelling-system.

Both semantics and phonology are considered when assigning a noun class. The Human/non-human distinction is an important feature. However, the last vowel of singular also seems to be an important feature for selecting the class form and agreement of the plural.

Tables are not used to report Kru noun class systems. Depending on the language, what will appear in the description can be either the NCA, in a non-alliterative agreement system, or the NCM. Gender is never discussed. Probably due to the fact that with a morpheme zero for singular, no pairing can be observed.

3.3. *Limits*

According to the literature and the analysis of the Excel table, what can be considered as the Kru noun classes framework does not use tables but paragraphs to explain which singular ending will trigger which plural form. Without table to summarize all the information about the noun class system, a typological use of this framework is merely impossible.
Another limit as a typological framework is the labelling-system. Either noun classes are not analyzed and qualified as such, or their labels will be too language specific.

Because Kru noun classes differ so much from other noun class systems found in Niger-Congo languages, the framework is also very different from Bantu or Atlantic noun classes frameworks.

However, a similarity is that each Kru noun class system framework encountered is too language-specific to be applied to other Niger-Congo languages. This language-specificity is amplified by the uncommon structure of Kru noun classes.

3.4. Other language subfamilies with similarities: KWA

KWA								
Tuwuli	2005	Noun classes	Number (Class/gender), letters (subclasses)	Phonology ; Semantics	NCF (NCM) ; NCA	Yes	Non alliterative	"Noun classes are determined by the plural prefixes alone, at least in those lasses which have an inflectional singular/plural distinction." Classes are considered as gender ; subclasses. NCA sometimes mandatory, sometimes not ? Tableaux p.104 et 107
Baoule	1977							By Creissels : pas des classes nominales mais classes nominales?????
Efutu	2008							There seems to be prefixal noun classes but the writer does not analyze them that way.
Tafi	2013	Noun classes	Morpheme	Phonology ; Semantics	NCF (NCM)	Yes		Numbers in the table but never used in the paper. Pairing (sg/pl) with links, but never used the term gender. "do not show agreement between nouns and their qualifiers."
Logba	2008	Noun classes	morpheme	Phonology	NCF(NCM) ; NCA	No	Non alliterative	Ne montre pas les liens entre NCM et NCA. "The nouns which have the o-/ɔ- verbal concord are by far the largest group of nouns. They comprise nouns with the following prefixes. o-/ɔ-, u-, e-/ɛ-, and a- prefix nouns18. These are all singular nouns. Nouns that trigger the i- verbal concord are those that take i- noun prefix. Nouns whose stems belong to the o-/ɔ- noun prefix take the o-/ɔ- sin- gular class. Nouns that are cross referenced by the N- prefix as verbal concord are those nouns that take the N-noun prefix. Nouns with the e-/ɛ- plural noun prefix trigger the concord of the same form." Focuses mostly on phonology, very rarely talks about the semantics. However it seems to be related : p. 70-71 (48-49)

Table 1.5: Kwa noun classes descriptions

According to Greenberg (1963), Kru languages are part of the Kwa family.

Figure 11: Greenberg classification (1963)

However, the study of many African languages has made it possible to refine the classification of Niger-Congo languages, which has led to changes in their classification. "Several researchers pointed out a greater relationship of certain Kru languages with Mande or Gur languages and subsequent research has led to the conclusion that it was autonomous. "(Platiel & Kaboré, 1998) . Thus, one of the major changes in Greenberg's classification concerns Kwa, which is now divided into four groups according to Williamson's classification (1989): "an independent family, New-Kwa, which corresponds to Greenberg's 'Western Kwa', while his 'Eastern Kwa' is split into two independent families: Kru and Ijoid, and a sub-family, Yroboid, which is integrated into the large Benue-Congo family (Platiel & Kaboré, 1998)."
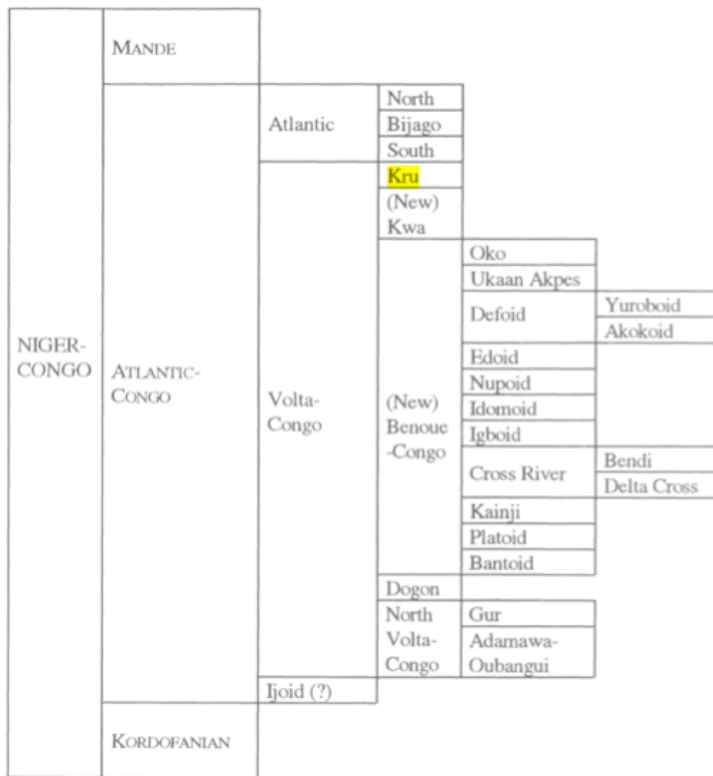
Figure 12: Williamson classification (1989)

Most of Kwa languages have lost their noun classes; they do, however, retain vestiges of them. According to Joseph: "Many of these languages have reduced their class system to the Animate vs. Non-Animate agreement pattern"[19] (Joseph, 2009, p.20).

Ega (Ivorian Kwa language enclosed in Kru territory) has preserved the complex system of noun classes like Bantu languages, just like Tuwuili or Gidere, also Kwa languages.

Let's not mistake here, the Kwa noun class system is very different from the Kru noun class system. If the Kwa noun classes are being discussed in this section it is because of the Tuwuli case, which is a Kwa language, but its noun class system shares the same features as the Kru noun classes: "Noun classes are determined by the plural prefixes alone, at least in those classes which have an inflectional singular/plural distinction." (Harley, 2005 ,p.88-89)

---

[19] "Beaucoup de ces langues ont réduit leur système de classes au schème d'accord Animé vs. Non Animé."

A big distinction with the Kru noun class description is the acknowledgement in each grammar of noun classes. Most of them will use tables to describe/report them. However, a similarity with the Kru noun class descriptions is the labelling, sometimes morphemes, sometimes a language specific framework (ex : Tuwuli). Phonology seems to also be an important feature for Kwa noun class description. When the system is alliterative, the NCA is always analyzed, which/this is the case in all the studied grammars in the Excel table.

"Tuwuli's noun-class system is of particular interest because it is motivated primarily by the plural forms rather than by the combination of the singular and plural forms, as is typically the case in Bantu languages" (Harley, 2005). No Kwa language has the same agglutinative richness.

Konoshenko, & Shavarina (2019) argue that "while nominal class systems in Kwa are rather simply classified as vestigial and fully functional or active, the degree of elaborateness in noun class systems of these languages is in fact quite diverse. On the one hand, languages with no active noun classification, e.g. Attié (Nyo), show traces of older class morphology. On the other hand, in languages with most elaborate nominal class sys- tems, e.g. Lelemi (Lelemic) or Tafi (Avatime-Nyangbo), various patterns of noun class system simplification and restructuring are attested. "

Harley (2005) describes in detail the complexity and richness of Tuwuli noun classes. According to him, "in Tuwuli, nouns are first divided into two groups, depending on whether or not they take both singular and plural prefixes, or just singular prefixes. Those in the former group are then classified according to their plural prefix alone, irrespective of their singular prefix. This double division yields six noun-classes in each group, since there are six different plural prefixes for nouns which have a singular/plural alternation, and six different singular prefixes for nouns which only have a singular form. The reason for categorizing nouns this way is that it preserves a degree of semantic unity within several of the noun-classes" (Harley, 2005, p.32-33)

With 22 different singular/plural pairs and six classes in the singular only, Tuwuli is the Kwa language with the most developed noun class system.

4. Gur Descriptions

The Gur family is the last Niger-Congo subfamily presented in the Second Chapter of this paper. Gur languages used to be called Voltaic languages. There are 85 Gur languages spoken by 20 million people. Gur languages are spoken in Southeastern Mali, Northern Côté d'Ivoire, Burkina Faso and Northern Ghana, Togo, and Benin.

Gur languages are divided in two groups: Central Gur and Senufo. Central Gur has more languages than the Senufo group which only counts 20 languages.

In many Gur languages, tonal systems are very important and tonal morphemes can be used to distinguish between grammatical functions. (Bendor-Samuel, 2020).

Noun classes in Gur languages are suffixal morphemes.

In this section, the description of Gur noun class systems in the literature will, first, be adressed, followed by the results of the analysis of the Table 1.7. We will then try to understand why such frameworks were applied and what are their limits for a typological framework for Niger-Congo noun classes.

4.1.   Description, literature

According to the literature, the Gur subfamily does not really possess a framework. Indeed, "up to this day each author working on Gur languages follows his or her own numbering system, agreeing only in the case of gender 1/2" (Miehe & Winkelmann, 2007, p.7). Miehe and Winkelmann also emphasized the importance to "hold that the convention of fixed numbering system for noun classes", which is considered "indispensable for comparative purposes" (Miehe & Winkelmann, 2007). In this regard, they proposed a common framework with a fixed numbering system for Gur noun classes. This framework is widely based on the Benue-Congo system. In the proposed Gur framework, each class has a number, and the pairing of some classes can form a gender, following the regular pairing (odd and even numbers). The Gur framework "follows the established Benue-Congo system as far as the numbers 1-19 are concerned" but will differ from it for class 23 in the Gur framework called 6a in the Benue-Congo for example.

The Gur noun classes framework, created by Miehe and Winkelmann in 2007, was not used for other descriptions. This framework was not the one found in most of the Gur languages descriptions studied for the Excel table. These descriptions were written between 2014 and 2018. Therefore, the literature about Gur noun classes seems to have had little impact on the noun class frameworks used for their description. However, one of the languages descriptions in the Excel table, written in 2018, does follow this framework.

In Senufo languages, the semantic is not as important as in Central Gur for the assignment of noun classes. The focus will mostly be made on "the agreement that nouns trigger on other elements associated with the nominal domain" (Traoré & Féry, 2018).

4.2. Excel Table

GUR								
Koulango	2016	Suffixes de classes nominales	morpheme	Phonology ; Semantics	NCF (NCM) ; NCA	No	Non alliterative	"schèmes d'accord" pour le défini, le qualificatif et le pronom, les différents genre sont animé, inanimé, mais c'est pas vraiment dit.
Supyire	1994	Noun genders	Number (gender)	Phonology ; Semantics	Gender ; NCM ; NCA	Yes	Non alliterative	
Tagbana	2018	Nominal classes	Number	(comments*)	Gender ; NCM ; NCA	Yes	Non alliterative	Number for classes and for genders. (4 genders, 7 classes) ; *Focus ni sur la sémantique ou sur la phonology mais sur l'agreement ... Syntaxe ?
Kusaal	2018	Noun class system	Number	Semantics	NCM	Yes (+links)		Pairing are established but never called gender. The word gender does not appear. There is no way to tell if it is an alliterative system or not.
Chakali	2017	Noun classes	Number	Phonology ; Semantics	NCM	Yes		"class can be regarded as phonological and/or semantic features encoded in the lexemes for the selection of the proper pair of singular and plural suffixes." Le NCA n'est pas discuté du tout on sait pas
Biali	2014	Système des classes nominales	Roman numerals	Semantics	Gender ; NCM ; NCA	Yes	Non alliterative	Roman numerals for gender, nothing for the classes

Table 1.6: Gur noun classes descriptions

As said above, the proposed framework by Miehe and Winkelmann (2007) is not encountered in most of the recent Gur languages descriptions.

Kusaal description is the only one following the Miehe &Winkelmann system: "In the present analysis, we synthesize the views on Kusaal noun classes into an elaborate system following Miehe *et al.* (2012) and adopting the conventions used therein (Musah, 2018)".

Most of the descriptions studied for the Table 1.6 refer to the noun class system as noun classes, except for the Supyire description in which noun classes will be analyzed through the scope of genders, called "noun genders" (Carlson, 1994).

Regarding the labelling system, most of the descriptions opted for a numbering system. Despite this similarity, all these systems are different. For the Supyire and the Biali descriptions, numbers will correspond to different genders. For the other ones, numbers correspond to different classes. The Koulango grammar only used morphemes to label noun classes. Labelling systems among Gur noun class systems are not unified.

Phonology for Gur noun classes descriptions seems to be an important feature compared to other Niger-Congo subfamilies. With or without a phonological aspect, Gur noun class systems seem to rely on semantics. However, in Tagbana, it was possible to observe a new focus. Indeed, in the Tagbana noun classes description, the author does not focus on phonology nor semantics but on the morphological dimension to assign a word to a noun class. This morphological assignment for gender is discussed in the Third Chapter of this paper (cf. Chapter 3, 3. Typological challenges).

All the Gur languages descriptions analyzed on the Table 1.6 use table to report on noun classes, except Koulango. In each of these tables, and in the Koulango noun classes description, the NCM always appear. The NCA also seems to be important for most of these descriptions. With no surprise, if a noun class system is non-alliterative, the NCA will appear in its description. The gender is not acknowledged in all of these descriptions but will never be found without NCA.

Regarding the alliterative feature, all the noun class systems are alliterative. It was not possible to establish the alliterativity or not for Kusaal and Chakali noun class systems because this matter was never addressed in their descriptions.

4.3.   Why this framework for this subfamily

After reviewing descriptions in the Excel table, there is no consistent or influential framework that emerge for Gur noun classes.

4.4.   Limits

The limits reside mostly in the diversity of frameworks than in a common framework itself. First because there is no common framework, but also because if we look at each framework individually, it is too language specific.
Mostly, the biggest limit for Gur noun classes description is the lack of consensus on the framework. Gender is not always appearing, however always treated. For example, in Koulango, it is implied that the two genders are animate and inanimate, but the author never uses the term gender nor overtly states it. The same limit is encountered in the Kusaal description. Each class is paired with another (or called a single class) to show the gender, but this pairing is never called a gender.
Another related limit is terminology. In the Tagbana grammar noun classes are called "Nominal classes" which usually refers to Nominal classifiers (cf. Chapter 1). This type of differences can easily lead to confusion and misunderstanding. Besides, a framework with a different terminology is hardly typological because it can be misunderstood by other research so the data can become obsolete. It is important to use transparent terms to reduce misunderstandings and oblivion as much as possible.

Conclusion

To study the different specificities of the classifications of Niger-Congo languages, we have created a matrix in an Excel spreadsheet. It allowed us to have a global view of each sub family of Niger-Congo languages, of their noun classifications when they existed, as well as of their characteristics and common points and differences. Moreover, the languages selected for this study were chosen according to available grammars. When there were several grammars, the choice was made on their relevance, quality, and recognition. The most recent grammars were prioritized. Each column of the matrix represents a characteristic necessary for a global vision of the descriptions of Niger-Congo noun classes. After having filled in

this matrix, we noticed that the different frameworks are sometimes so different and so specific to each studied language that it is difficult to extract a common framework. Indeed, even if the concept of noun classes is common to each language, which is the main characteristic of Niger-Congo languages, the way of describing, classifying, and analyzing them is different from one language to another. In the case of the noun classes of the Bantu languages, based on the proto-Bantu system of noun classes, the classification consists in assigning a class number to the nouns, ranging from 1 to 20, sometimes adding letters. On the other hand, in the case of Atlantic languages, noun class systems, when they exist, generally refer to noun classes using morphemes and there is no coherence in the frameworks used by linguists, mainly because of the lack of proto-Atlantic reconstruction. In the case of Gur languages, noun classes are suffix morphemes. We note that there is no common consensus on the classification of Gur noun classes and that each linguist works with his own numbering system. Thus, there is no consistent or influential framework for Gur noun classes. As for the Kru languages, their noun class system is nowadays almost non-existent, and most linguists do not even use the term noun classes when referring to Kru languages. The few class markers that still exist is the notion of human. To proceed in the elaboration of a typological framework, generalizable to all Niger Congo languages, it will be necessary to mix all the characteristics to make up one framework, less specific and adaptable to all languages.

CHAPTER 3: Challenges

Introduction

Chapter 2 has highlighted the issues to be met to set up a typological framework, common to all noun classes of African languages. Indeed, while detailing the descriptions of noun classes of the families of Niger-Congo languages, many challenges appeared, such as the lack of unity on how to deal with noun classes, but also the diversity of noun class systems or, simply, the lack of data for some little studied languages. This third chapter will therefore be devoted to highlighting, describing and analyzing the challenges faced when linguists attempt to create a typological framework for noun classes. It will answer the question: What are the challenges in creating a typological framework for Niger-Congo noun classes?

To succeed in creating a common typological framework for noun classes, it will be necessary to address various challenges that have been divided into three main categories:
- finding common ground to deal with ever-changing noun classes,
- the different languages specificities,
- finding a way to match the genders across languages,

Obviously, for reasons of time and space, the list of challenges should be considered by the reader as non-exhaustive. Only the main challenges that need to be addressed and dealt with are listed here. As we will see, these challenges are all intertwined, one leading to the other.

To list the challenges chronologically, we asked ourselves the question: what different stages of challenges do we encounter when attempting to create a typological framework for Niger-Congo noun classes?

Before studying the different challenges, it seems essential to define what a typological framework is. Typology is the categorization of phenomena into various types. For linguists, typology is essential because it allows to "determine the most fundamental properties of language" (Sapir, 1907, p.34). As R. Jakobson puts it: "the study of patterns that occur across languages, with the aim of formulating typological generalizations about patterns". While for Greenberg, "linguistic typology is an approach to linguistic theory that uncovers and attempts to determine the limits of human language by looking for (near-) universal patterns in language structure and to explain them in terms of linguistic function."

A typological framework would allow linguists to make statistics and find out intra-genetic statistical and implicational universals for noun classes, the absolute intra-genetic universals being part of the framework. In the case of noun classes, with a typological framework it would be possible to define the different "universals" among Niger-Congo noun classes, also known as intra-genetics universals. What would be considered as universal actually is not universal because it only applies for Niger-Congo noun classes, not every language of the world. These universals will only concern a portion of the world languages: Niger-Congo languages with productive noun class systems. Productive in the sense that they show existing agreement patterns, and noun classes are not only traces in the language but still a functional linguistic tool.

From this definition and this contextualization, we can better identify the different difficulties encountered and the different challenges that will have to be met to create a typological framework for Niger-Congo noun classes.

1. No common ground

1.1. The analysis is constantly being revised

The first and greatest challenge is that analysis and theories around languages are in constant evolution, and therefore always under revision, from one study to another, from one linguist to another.

A big challenge in the creation of a typological framework for noun classes lie in the unstable theories surrounding noun classes. What used to be true about noun classes is not anymore and vice versa. In a very natural evolution of science, each theory is building on the other to always go further, so when are we sure we are holding the right one and should base our framework on it? In this section, two compelling examples of different theories or analysis of noun class systems will be presented. The first example shows how from one theory a new one can be born, but also how two theories, even if one is inspired by the other, can greatly differ from each other. The second example, shows how, by changing the framework, or the way we analyze a noun class system, we can have an impact on the data.

In the first instance, Güldemann (2019) draws on a theory of gender and agreement developed by Corbett to reach a different conclusion. Corbett's (1991; 2000; 2006) work has served as the primary reference point for the previous typological analyses of gender and related phenomena. Güldemann (2019, p.97) applies "a strict distinction of four concepts, which are necessary whenever gender is reflected by syntactic agreement as well as nominal morpho-phonology, the latter implying

some amount of what Corbett (1991) calls formal class assignment. An agreement class in the present conceptualization is thus a set of noun forms that share an identical behavior across all agreement contexts of a given system and thus equals what Corbett (1991, 2006) calls a "consistent agreement pattern"" (Güldemann & Fiedler, 2019, p.101These two approaches have different extensions and concerns, yet they partly overlap which can create confusion.). Our framework also departs, in some respects, from Corbett and Güldemann (cf. Chapter 4, 2. Discussion).

In the second instance, Wolof noun class system descriptions offer a good example of the perpetual evolution of theories and analyses around a language, especially in the analysis of noun classes. These changes can be rapid. Thus, between two grammars published one year apart in 2015 and 2016, theories are opposed, and we can see a different approach to noun classes, which leads to fundamental changes such as the number of classes. Thus, in 2015, Pozdniakov & Robert identified 10 noun classes in Wolof (K, B W, M, G, J, L, S, N, Y) 8 singulars and 2 plurals. While Babou & Loporcaro, (2016) consider that two other plural noun classes exist and according to their analysis (abstract): Wolof turns out to have a complex gender system, featuring 17 distinct gender values.

Thus, in Noun classes and grammatical gender in Wolof by Babou & Loporcaro, (2016, p.2): "By making use of these notions, we will show that our data reveal interesting differences from the currently available descriptions of aspects of the morphology and morphosyntax of noun classes in Wolof. These aspects have not been satisfactorily elucidated so far, partly because most analyses have focused exclusively on (2a)[=Noun classes] (defined in terms of word forms), without paying sufficient attention to (2b)[=Agreement classes or genders], i.e. the agreement pattern chosen by the nominal lexeme as a whole."

1.2. Non-reconstruction of Proto Niger-Congo Noun Class System

The last challenge regarding the lack of a common ground is due to the fact that there is no reconstruction of the Proto-Niger-Congo noun class system. According to Creissels and Pozdniakov (2015), it "is useless to systematically compare Atlantic class markers without an internal reconstruction for each system, so that analogic changes can be updated"[20].

This reconstruction could allow linguist to create a global framework the same way Bantuists created a framework following Proto-Bantu noun class systems. This consistency would be a major achievement for a typological approach of Niger-

---

[20] "[Il] ne sert à rien de comparer systématiquement les marqueurs de classe atlantique sans avoir procédé au préalable à une reconstruction interne dans chacun des systèmes, de façon à mettre à jour les changements par analogie. "

Congo noun classes. Noun classes being an intra-genetic phenomenon, it can be helpful to include genealogical information to their typology, though it is not mandatory.

This reconstruction has been a challenge for historical linguistics. Indeed, the genealogical classification of African languages is still considered as a controversial question. The Niger-Congo family is especially criticized due to its diversity.

2. Language specificities

2.1. Marking

A challenge when attempting to create an encompassing framework for Niger-Congo noun classes is the different forms a noun class can take depending on the language.

Some languages will use tones, whereas others will use a morpheme zero. Another difference can be found in Atlantic languages with the stem-alternation. Sometimes, a noun class can undergo phonological processes and conditions which can impact its form. The marking of noun classes is a primary challenge.

Tonal noun classes are a serious difficulty to report. Probably because, usually, in a language, when noun classes use tonal distinction, it will not be applied on the entire noun class system. So, linguists need to mix in a same framework, the reporting of morphemes and the reporting of tonal morphemes. Nchane, a Bantoid language, displays this type of behavior. "The singular/plural distinction of gender 9/10 nouns is indicated by tone alone […] The class markings consist of floating tone prefixes in both classes, a floating L for class 9 and a floating H for class 10" (Boutwell, 2020). However, the author reminds us to take this analysis cautiously. Indeed, it seems that this phenomenon was a challenge, first to understand, second to report. He decided to report the tonal morpheme among the other morphemes by adding a tone on a σ.

Class	Nominal marking	Agreement marking[31]	Class	Nominal marking	Agreement marking[31]
1	Ø-	wu	2	bā-	ba
3	-w-, Ø-	wu/kfu	4	Ø-	yi/che
5	Ø-	chi	6	ā-	a/ka
7	kī-	ki	8	bī-	bi
9	σ̀-	yi	10	σ́-	yi/che
14	b(v)ū-	b(v)u	6a	māN-, N-	maN/ma/mu
19	fī-/fīN-	fi	18a	mū-/mūN-	muN/mu
			13	chī-	chi/yi
16	fɛ̄-	fɛ/fɔ			
18	ā-	a/ya			

Table 5.1 Summary of Nchane noun classes.

Figure 13: Nchane noun class system – (Boutwell, 2020)

A compelling example of languages using the morpheme zero for noun classes is the Kru family. As discussed previously (cf. Chapter 2, 3. Kru descriptions), the singular classes are not overtly marked, thus considered as a morpheme zero. The morpheme zero is a challenge in the sense that it makes it more complicated for linguists to detect noun class pairings. However, it is already widely accepted that a morpheme zero needs to be marked that way: Ø.

The last example of language-specific challenges is the stem-initial alternation found in North Atlantic languages. Because in most of these languages, "marking on nouns and class agreement crucially involves not only affixes, but also stem-initial alternations" (Creissels, 2015, *p.16*), the stem-alternation phenomenon needs to be considered as part of the noun class system. Moreover, stem-initial alternations contribute "to the expression of class distinctions in Atlantic languages" (Creissels, 2015, p.16).

Creissels adds "however, they are morphologized to a considerable extent, and the analysis of the underlying processes (be it in synchronic or diachronic perspective) is not an easy task", emphasizing the difficulty lying in the reporting of stem-alternations in a typological framework.

It is important for a typological framework to take this phenomenon into account and to give a clear picture of it.

A way to report tones, as much as stem-alternation for example, is crucial in the creation of a typological framework able to gather data from all Niger-Congo languages with productive noun-classes.

2.2. *Derivational Use*

As Creissels explains (2014, p. 571), derivation is a very common phenomenon in Niger Congo languages and can be found in different forms: "In typical Niger-Congo languages, verb-to-noun derivation typically involves two elements: the addition of a derivational suffix, and the addition of class morphology manifesting the assignment of the derived noun to a particular class, as in Jóola Banjal -ffaŋ'close' → e-ffáŋ-um (pl.si-ffáŋ-um) 'key', where -úm is a derivational suffix used to derive nouns from verbs, and e-/si- are class markers. It may also happen that no overt derivational element is present, and the deverbal noun is formed by the mere addition of class morphology to a verb stem, as in Jóola Banjal -mbal 'fish (V)' → e-mbal (pl. si-mbal) 'fish-trap'. In such cases, the class to which such nouns are assigned may be crucial for the identification of their meaning."

Still, the derivational use of noun classes may differ in many ways, depending on the language. To be helpful, a typological framework must take these differences into account.

Addressing the derivational use challenge entails to discuss inherent or non-inherent members of a gender. Some genders only have inherent members, these genders do not have a derivational use. On the other hand, a gender with derivational use can have inherent members, but not only. It will have non-inherent members.

Moreover, connotations to gender, which can also be used as derivations, are very much tied to culture and social context. These are very language-specific features of noun classes that need extra care when creating a typological framework.

In Bantu languages, "Some classes do not have any ''inherent'' members and are, instead, dedicated to specific semantic purposes. Class 20, for example, is used as an augmentative class in many Bantu languages, and Classes 12 and 13 are often diminutive, with few or no inherent members". (Morrison, 2018, p. 3)

Bena is a good example of this statement, "classes 12-13 have no inherent members and are used exclusively for diminutives, and Class 20 is used exclusively for augmentatives" (Morrison, 2018, p. 3). From a table showing the "semantic tendencies within Bena noun classes", and some examples, we can deduce that classes 5/6 and 7/8 are not considered as derivational because they have inherent members, as opposed to classes 12/13 and 20 displaying an only derivational purpose. The term augmentative is used in both cases, but the author shows a major difference between them.

Table 3
Semantic tendencies within Bena noun classes.

Class	Entities
1	Humans
2	Plurals of Class 1
3	Body parts, plants and their products, some landscape terms, other
4	Plurals of Class 3
5	Food, body parts, animals (usually medium- or large-sized), insects, some plants, some landscape and weather terms, singular augmentatives, other
6	Default plural for Class 5, plural augmentatives, sometimes used as plural for Classes 9 and 11
7	Animals, some diminutives (though not as small as Class 12/13), objects, a few body parts, some borrowings (especially from Swahili), other
8	Plurals of Class 7
9	Crops (collective-type plants), abstract nouns, objects, some animals, some body parts, geographic features, borrowings, other
10	Plurals of Classes 9 and 11
11	Some collectives, objects, other
12	Diminutives only (no inherent members)
13	Plurals of Class 12
14	Collectives, mass nouns
15	Verbal infinitives
16	Location (indefinite)
17	Location (definite)
18	Location (inside)
20	Augmentatives (no inherent members)

Figure 14: Bena noun classes semantic tendencies – (Morrison, 2018)

Another example of derivational use with noun classes can be found in Tswana. Though this language does not use diminutive and augmentative classes anymore, Tswana noun class system demonstrates an intensive derivational use (Creissels, 2014). This is illustrated in the following example showing that a noun belongs to different classes, according to what the speaker wants to express (Creissels, 2014).

a.	cl. 1	*mosadi*	'woman'
	cl. 2	*basadi*	'women'
	cl. 7	*sesadi*	'feminine behavior'
	cl. 9	*tshadi*	'group of women'
	cl. 11	*losadi*	'group of women'
	cl. 14	*bosadi*	'womanhood'
b.	cl. 3	*moretlwa*	'tree of the sp. moretlwa'
	cl. 2	*meretlwa*	'trees of the sp. moretlwa'
	cl. 9	*thetlwa*	'fruit of the moretlwa tree'
	cl. 10	*dithetlwa*	'fruits of the moretlwa tree'
	cl. 11	*loretlwa*	'thicket of moretlwa trees'

Figure 15: Tswana noun class system and derivation – (Creissels, 2014)

It this case, gender does not have inherent members but inherent meaning that can be applied to different nouns. This is considered as a derivational use and needs to be considered in a typological framework.

In Supyire, a Gur language with fewer genders than Bena, Gender 2 and Gender 3 are respectively named "Augmentative" and "Diminutive" by Carlson (1994):
- Gender 1: Human
- Gender 2: Augmentative
- Gender 3: Diminutive
- Gender 4: Collectives
- Gender 5: Pourable

Contrary to Bena, augmentative and diminutive genders have inherent members, such as 'big things' for Gender 2 and 'small things' for Gender 3 and display a derivational use too. Moreover, "moving a root into gender 2 may have pejorative force. Body parts normally in gender 3 acquire the added meaning of 'big an ugly' when put into gender 2. […] Similarly, nouns in gender 1, referring to human beings may gain a certain loutishness." Gender 2 and 3 display a derivational use as much as possess inherent members.

This shows that Niger-Congo languages use different strategies when it concerns the derivational use. These different strategies need to be understood and described in a typological framework for Niger-Congo noun classes.

2.3. Loan Words

Loan words are undoubtedly to take into account in the creation of a typological framework because their behavior in noun class systems differs widely from one language to another. A gender loan words is not possible; however, it is the strategy for some languages. Gathering in the same framework these different strategies is a challenge, mostly when this framework also needs to be readable.
All the examples illustrating these different strategies are from *A comparative Study of Bantu Noun classes* (Maho, 1999), which explain why only Bantu languages will exemplify this challenge. However, it is even more to be considered as a challenge if differences can be observed within the same sub-family. We can speculate that wider or similar differences can be found in the rest of Niger-Congo languages. Maho's section on loan words is largely based on the work of Okhotina (1975).

The first strategy is very close to the meaning. In Gogo, "each loan word is included into a certain class according to its semantic and takes the prefix of the class, the formal patterns of class agreement". The Gogo noun class system is considered as "representing 'the most conservative' stage" (Okhotina, 1975 ; Maho, 1999).

The second strategy, found in Isizulu, a language with a noun class system displaying a 'more advanced' stage, is when though semantic is important for noun classes attrition, it had no impact regarding loan words. Indeed, in Isizulu, "a definite subset of classes is chosen to assimilate the loans irrespective of their semantics" (Okhotina, 1975 ; Maho, 1999).

The third strategy illustrated here can be found in Kiswahili and Lingala. It is considered as the "highest degree of evolution". Loan words will be assigned their noun class depending on the feature animate-inanimate (Okhotina, 1975 ; Maho, 1999).

It is necessary for a typological framework to reflect these three strategies, or more. Loan words are a challenge in the creation of a typological framework.

3. Challenges at a typological level

In this section, we will see that establishing 'typological cross-linguistic genders', in other words, corresponding 'synchronic genders' between Niger-Congo languages, is a huge challenge. These "typological cross-linguistic genders" can give the possibility for further typological analyses of Niger-Congo noun class systems.

First, the noun classes evolution and reduction are found in most Niger-Congo languages. In Bantu languages for example, noun classes are reducing to only distinguish between animate and inanimate (Maho, 1999; Okhotina, 1975). "Some structural changes in bantu languages are due to their specific communicative function". Damanan N'dré (2016, p.69) explains noun class system disintegration in Kru languages by "the irregularity of the forms of the plural suffix".[21]

For some languages, "the semantic criterion of class identification is [still] very operative" (Maho, 1999; Okhotina, 1975). However, as seen previously with Kru languages, it is not always the case.

The noun classes reduction phenomenon is to take very seriously because it will have a concrete impact on data. The reduction slowly makes noun class systems less overt in the language and less productive. Languages are all at a different stage of reduction, some Niger-Congo languages have already lost their noun class system productivity and only traces remain, in Yoruba for example. While some other languages, like Gogo, are considered very "conservative".

Which leads us to this typological challenge: how to make cross-linguistic genders correspond when they are all evolving in different ways and noun class systems are slowly being less and less productive in languages? On what should we base these "cross-linguistic genders"?

Nowadays, Niger-Congo languages use different strategies to assign a gender to a noun. Many will keep a semantic basis, but some will use phonological rules to assign a gender, like Kru languages for example. In Wolof, a very interesting gender assignment basis is morphological.

According to McLaughlin (1997), in Wolof, derived nouns will have a specific gender assigned. There are three derivational processes that can trigger this assignment: stem-initial consonant mutation, suffixation and reduplication. They "entail the assignment of the nouns to a specific noun class. For example, the manner suffix, /-in/, assigns a noun to the w-class, and those involving consonant mutation are generally assigned to the g-class. In such cases, the derivational affix […] assigns that class to the noun." (McLaughlin, 1997, p. 20). This type of assignment is a morphological assignment

Corbett and Fraser (2000) established a gender assignment typology stating that there are four different gender assignment systems among languages of the world. In this paper, only three of them will be used: semantic assignment, morphological assignment, phonological assignment. The semantic assignment is self-explanatory and is the most common among Niger-Congo noun class systems. As seen previously, Kru languages predominantly assign gender using phonological rules.

---

[21] « L'irrégularité des formes du suffixe de pluriel, qui ne tient nullement compte d'une analyse phonologique solide est certainement l'une des causes majeures de la désagrégation du système des classes nominales dans les langues kru ».

This morphological assignment system corresponds to the example from Wolof described above.

We can see that Niger-Congo noun class systems gather many different gender assignment systems which can lead to many complications.

Not sharing the same basis for assignment is a challenge because we have nothing in common between all Niger-Congo noun class systems. How is it possible to create a common basis when it does not exist? We wish we could have typological cross-linguistic 'genders', but because their assignment is due to different reasons, it is hard to create a common pool. Therefore, we can create an encompassing typology for these variations between noun class systems which need to be measurable. It will be important for a typological framework to take into account the different factors in gender assignments in order to capture these differences within the same frame.

To conclude this section, it is very complicated to have typological cross-linguistic corresponding 'genders' for Niger-Congo noun class systems because they do not share a common gender assignment basis.

Conclusion

When one attempts to create a typological framework for Niger-Congo noun classes, four major groups of challenges emerge chronologically. The first one consists in finding a common ground to deal with noun classes whose descriptions are in constant evolution, as for example the description of Wolof which gained two classes in the interval of one year and two different linguistic studies.Getting linguists to move towards a new framework can also be a major challenge.The second challenge is to bring together, in a single framework, the linguistic specificities of the different Niger-Congo noun class systems. The third challenge is to find a method for matching genders between languages, since even though many languages rely on semantics to assign gender to a noun, others use phonological rules such as the Kru languages or morphological rules such as Wolof. For this framework to be truly typological, it must be filled with as much data as possible and, needless to say, with reliable data that is as representative as possible of the noun class systems of Niger-Congo languages.

CHAPTER 4: Criteria and Proposed Framework

Introduction

This chapter is the outcome of the project to create a common framework and database for Niger-Congo noun class systems and their implementation. This idea was born from the observation of a lack of such a common framework and from the desire to fill it.

Previous chapters have contributed to the elaboration of this framework. We have reviewed the definitions of noun classes and the reasons why a typological framework seems important and the challenges that creating such a framework would entail. This fourth and final chapter is devoted to a description of this framework, its analysis, and limitations.

First, we will review the criteria and characteristics that are essential to a typological framework for Niger-Congo noun classes. Then, we will describe in detail the proposed framework and database, their advantages, and their usefulness both for linguistic documentation and for typological analysis. In the second part, solutions to some of the challenges mentioned above and highlight the advantages of such a common framework will be proposed. We will discuss the objectives and the advantages of this framework, allowing to do predictive analysis and statistics. For the sake of intellectual rigor, the limitations of this framework will be also discussed, together with some solutions to remedy them.

1. Criteria and Main Features

1.1 Criteria for a Typological Framework

For a typological framework, common to all Niger-Congo noun class systems, to emerge, 3 major criteria must be met: encompassment, flexibility, and predictability.
Indeed, challenges stated previously actually refer to the diversity of structures found in Niger-Congo noun class systems, in spite of a common trunk which defines them as noun classes. A typological framework needs to be flexible and encompassing. However, it needs to keep some rigidity in the sense that it needs to be consistent and gather relevant information for a typological analysis. A flexible and encompassing framework, if strict enough, will be clear and

readable. The goal of this framework is to enable linguists to spot similarities and differences between Niger-Congo noun classes in an accessible and readable manner. Indeed, "Linguistic typology is interested in both differences and similarities between the languages, because these are interrelated" (Krasnoukhova, 2019).[22]

1.1.1 Encompassment

Encompassment is the capacity to enclose and to envelop. For a framework, being encompassing consists in gathering a diversity of data. Niger-Congo noun classes display a high diversity, as much in the way they behave as in the way they are described. Although, this typology of Niger-Congo noun classes only focuses on the data, not their descriptions. An encompassing framework is a framework allowing linguists to apply it to all Niger-Congo noun class systems. The framework can be applied to data from a Bantu language as much as from a Kru language, though these data display different features. If the framework is encompassing enough, it is possible to consider the diversity found in Niger-Congo noun classes.

A common typological analysis is only possible if the frame is sufficiently generic to be applicable to every Niger-Congo noun class system. The frame still needs to be detailed enough to capture differences between languages.

1.1.2 Flexibility

One is considered flexible when able to change or to be adjusted to cope with variable circumstances and to meet particular or varied needs. In other words, flexible implies closely related but slightly different notions: first, the ability to change to overcome unexpected challenges, second, the capacity of being changed or adjusted to address specific needs or questions.
A flexible framework must give the opportunity to be changed, if necessary. For example, if new data went to reveal a new feature, it should be possible to consider it. It will also need to enable researchers to specify data or questions. For instance, if a researcher is interested in a specific feature in Niger-Congo noun classes, a flexible framework should enable them to adjust their frame to make data more accessible and readable.

---

[22] For more references on linguistic typology, Whaley (1997) and Velupillai (2012) introduce the topic. Krasnoukhova's (2019) introduction to typology is a recent introduction to linguistic typology.

A common typological analysis is only possible if the framework can adapt to evolving data and research questions.

1.1.3 Predictability

To meet specific typological analysis needs, the framework needs to offer the possibility to develop statistics from available data. By providing an access to statistical analysis, the framework can help to predict common behavior. The predictability dimension can be used on two different levels.

The first level is intra-linguistic. Within the same language, the framework can display predictability. It means that, based on available data, it should be possible to determine to which gender a word belongs. For example, if we encounter a singular sentence, we should be able to predict the plural forms words should take because we will be able to know in what gender this noun belongs and what the NCA is. This type of prediction becomes impossible in case of homophony between noun classes, which is why not only the NCM needs to be taken into account. A framework showing the gender by using the NCM and the NCA, among other important features, offers reliable predictability. "The notion of class is predictive of other paradigms cells" (Hepburn-Gray, 2018, p.32).

The second level of predictability resides in the typological, thus cross-linguistic, analysis. To meet typological analysis needs, such as intra-genetic statistical or implicational universals, the framework needs to offer the possibility to develop statistics from available data. By providing an access to statistical analysis, the framework can help to predict common behavior between languages. The framework should enable linguist to identify statistical intra-genetic universals within Niger-Congo noun class systems. It could also reveal absolute intra-genetic universals still unknown until now – still, it is risky to call them universals due to the fact they are not common to all languages in the world and only apply to Niger-Congo noun classes.
With the ability to develop statistics, it would be possible to answer a question such as "Are suffixal noun classes more inclined to have a morpheme zero as a singular NCM? ". In order to have a predictive dimension, the framework needs to offer the possibility to calculate correlations.

A common typological analysis is enhanced if the framework offers the possibility of being predictive using data in a statistical form.

1.2 Noun Classes Main Features for a Typological Framework

In order for the framework to be "encompassing", "flexible" and "predictive", 8 main features need to appear in the table. These features correspond to the basis of Niger-Congo noun class systems structure. They represent the core of noun classes and make the framework encompassing and flexible. These features, discussed below, are labeled: Gender/NCM-cum-NCA; NCM; NCA; Affixation; Assignment; Meaning; Derivational use; and Alliterative.

Gender / NCM-cum-NCA
Gender is supposed to be specified. However, as seen in Chapter 3, it is part of the typological challenges. It is complex to assign a name or a number to a gender without being too language-specific or arbitrary. A solution to this difficulty is proposed in the second part of this chapter. The gender/NCM-cum-NCA will enclose the NCM and the NCA triggered by NCM as defined by Hepburn-Gray (2018). This "gender" is different from the one proposed by Corbett (1991) or Güldemann (2019) because it takes into account the NCM. There will be more NCM-cum-NCA than genders as defined by Corbett. An NCM-cum-NCA can consist of a SG-PL pairing or a single class.

NCM
The Noun Class Morpheme (NCM) is the form taken by the noun class on the noun. The singular and the plural forms for each gender/NCM-cum-NCA will need to appear in the table. In case of a single class gender, the NCM will still need to be specified. Allomorphs and their conditioning will also be stated.

NCA
The noun class agreement pattern (NCA) needs to appear too. It consists in giving the form of each triggered noun class morpheme for each concordial element. This is because the dependent elements can sometimes be marked differently from the noun. These different forms need to appear, showing to which target element they correspond to (Adjective, Possessive, ...).
Noun classes can be marked as subject and/or object on the verb. It is important to display these forms, as subject and as object.

Affixation
The type of affixation needs to be specified. Niger-Congo noun classes affixes can be prefixes, suffixes, infixes, and circumfixes (cf. Chapter 1, 2. Different descriptions and definitions). Generally, they are prefixes or suffixes, however, few can be infixes and circumfixes. Though the frequency

of each is very unequal, it is important for all these possibilities to appear in the table for the typological framework to be encompassing.

Assignment

The gender is assigned to a noun for different reasons. Corbett (2000), in his typology of gender assignment, differentiates four different systems among languages:

- Semantic assignment,
- Predominantly semantic assignment,
- Morphological assignment,
- Phonological assignment.

Interestingly, Niger-Congo genders formed by noun classes can relate to all of them. Sometimes, different systems can be found among the same language, in some Kru languages for examples, which mix semantic and phonological assignment systems (cf. Chapter 2, Kru descriptions). Only the three assignment systems (semantic, morphological, and phonological) need to appear in a typological framework for Niger-Congo noun classes with the possibility to mix them. Indeed, sometimes linguists will encounter noun class systems in which the assignment is mixed.

As seen in Chapter 3, it is important to specify what is the reason why a noun belongs to a gender. The framework should show the assignment for each gender.

Meaning

When the gender assignment is based on the meaning, which is generally the case but not always, the possibility to give the meaning of a gender/NCM-cum-NCA must appear in the table. To remain consistent and typological, it is a good solution to follow Contini-Morava's (2002) tags: "The database tags are not meant to be mutually exclusive; more than one can be entered for a given noun. They were chosen based in part on my own knowledge of Swahili language and culture, in part on categories found relevant in earlier cross-linguistic work on noun categorization (e.g., Adams and Conklin 1973; Craig 1986), and in part on categories recognized by earlier studies of Swahili and other Bantu languages (e.g Meinhof 1948 [1906]; Ashton 1944; Polomé 1967; Denny and Creider 1976; Hinnebusch 1979; Zawawi 1979; Spitulnik 1987, 1989). The categories are avowedly "etic": they were used in order to make it possible to manipulate large quantities of data with reasonable flexibility, but they do not constitute an analysis. And in fact the analysis does not reflect the database tags in a direct way" (Contini-Morava, 2002, p.53).

Derivational use

It needs to be possible to discuss the derivational use of each gender. As seen in Chapter 3 (cf. 2. Language specificities, derivational use), some noun class systems approach the derivational dimension differently.

We will consider three possibilities in the derivational use for Niger-Congo noun classes:

- First possibility, the gender is only used for derivational purpose.
- Second possibility, the gender can have a derivational use, but also has inherent members.
- Third and last possibility, the gender only possesses inherent members, thus is never used for derivational purpose.

Alliterative

When the NCM and the NCA are similar, which means they share the same form to the extent of some allomorphs, if they are conditioned by the general phonological rules of the language, the noun class will be considered alliterative. On the other hand, if the NCA, triggered by the gender, does not share the same form as the NCM and the allomorphs are different due to a morphophonological reason, the system will be considered non-alliterative. The possibility to mark the gender as alliterative or not makes the framework more flexible and readable. We will not have to fill the NCA section if it is specified that the system is alliterative.

2. Discussion: Proposed Framework and Database

In this section, I design solutions to some of the challenges discussed previously. It occurred to me that the creation of a database, a kind of prototype, would serve as a basis for a typological framework for Niger-Congo noun classes. Of course, the following explanations should be taken carefully by the reader, who needs to approach them with a critical eye in order to improve this proposition.

2.1. Why a Database

A complex challenge to overcome when attempting to create the framework for Niger-Congo noun classes is the creation of typological and non-arbitrary labels for genders/NCM-cum-NCA (cf. Chapter 3, 3. Typological challenges). I believe it will be easier to use numbers to identify each gender. However, they will not correspond to already known labelling-numbers such as the numbers for Bantu noun classes.

A database management tool enabling to pool as many data as possible proves to be a powerful way to assign a consistent and typological number to each gender/NCM-cum-NCA.

In this database, linguists will be able to insert the mandatory features listed in the first section of this Chapter. So, the NCM with their corresponding NCA and meaning will need to be filled and will constitute a "gender", which will be referred to as NCM-cum-NCA (cf. Chapter 4, 1.2. Main features) to prevent confusion with Corbett's gender which only focuses on the NCA. For the meaning and the framework to stay consistent, linguists will have to select pre-made tags that will give a very general meaning, they will be able to create new tags if necessary, and accumulate tags for a same gender/NCM-cum-NCA. Pre-made tags will consist in very basic meanings that needs elaboration, we will find for example: "Human", "Animate", "Inanimate", "Collective", "Loan words", "Tree", "Animal", "etc. Once there are enough data to analyze, we should be able to use this database to obtain typologically synchronic applicable numbers for each gender/NCM-cum-NCA found in Niger-Congo languages based on their functional and structural similarities. It means that each NCM-cum-NCA should correspond cross-linguistically. Some will not appear in some languages because they do not possess it.

2.2. Framework and Database Presentation

At the present stage, the database prototype was developed on an Excel spreadsheet but, if necessary, it could be developed using another, more appropriate, software specializing in relational data management systems such as dBASE or SQL SERVER. The same fields in the database must be filled in for all the languages studied. The data of any Niger-Congo noun class system should fit into this table.

The Table consists of a pair of tables dependent of each other: the main table and the table "To fill". Here is an example of what each individual table would look like (cf. Excel document: Typological framework proposition Chapter 4).

Main Table (reference values)

		Affixation	Pairing			Morphophonological rules	Yes		Yes	Semantic	Semantic		Yes only	Primary
		Prefix	Pairing			Morphophonological rules	Yes		Yes	Semantic	Semantic		Yes only	Primary
		Suffix	Single class			General phonological rules	No		No	Morphological	Phonological		Yes not only	Secondary
		Infix				No allomorphy					Morphological		No	
		Circumfix									Mixed			

Identification	Gender / NCM-cum-NCA	Affixation	Noun Class Morpheme (NCM)	Number of allomorphs SG	Number of allomorphs PL	Conditionning	Definiteness	Alliterative (NCA)	NCA agreement	Assignment		Derivational use	Hierarchy
#1		Prefix / Pairing	SG: m PL: wa	1	0	General phonological rules	Yes	SG: m PL: wa — No	See columns NCA	Morphological / Semantic	Human	Yes not only	Primary
#2		Single class	ku			General phonological rules	No	Yes	ku	Semantic / Semantic	0		Primary
#3			0			General phonological rules	Yes	SG: PL: — No	See columns NCA	Mixed			Primary
#4			0					Yes	0	Morphological	0		Secondary
#5			0					Yes	0	Mixed			Primary
#6			0					Yes	0	Semantic	0		Primary
#7			0					Yes	0	Semantic	0		Primary
#8			0					Yes	0	Phonological	0		Secondary
#9			0					Yes	0	Phonological	0		Primary
#10			0					Yes	0	Semantic	0		Secondary
#11			0					Yes	0	Semantic	0		Primary
#12			0					Yes	0	Phonological	0		Secondary
#13			0					Yes	0	Morphological	0		Primary
#14			0					Yes	0	Phonological	0		Secondary
#15			0					Yes	0	Semantic	0		Primary

Table 2.1: Main Table of the proposed typological framework

NCM		Allomorphs			definite NCM		NCA SG							NCA PL							Assignment		
		Forms					NP					VP		NP					VP				
SG / Single class	PL	SG / Single class	PL	Morphophonological Rules	SG / Single class	PL	Adjective	Enumerative	Demonstrative	Quantitative	Possessive	Subject	Object	Adjective	Enumerative	Demonstrative	Quantitative	Possessive	Subject	Object	Meaning	Phonological rule	Syntactic rule
m	wa	mw			m	wa					a	yu							wa	wa	Human		
ku					ku																		

Table 2.2: Table "to fill" of the proposed typological framework

Each table of the database will be presented according to this Excel spreadsheet, which corresponds to my proposition of a typological framework.

This table considers a gender/NCM-cum-NCA as the combination of a SG-PL pairing (or single class) with an NCA and, when possible, involving semantics or other features. Links as seen in Corbett or Guldemann do not exist with/in this framework to make it applicable to every Niger-Congo noun class system. Indeed, when a SG-PL pairing has the same form but a different NCA, it takes a new row and becomes a new gender/NCM-cum-NCA.

Example of the links by Guldemann (cf. Chapter 1, 4. Definition relevant for a typological approach of noun classes):

Figure 6: Gender system (left) vs. deriflection system (right) of Ikaan (based on Borchardt 2011: 75–78)

Figure 16: Gender system and Deriflection system of Ikaan – (Güldemann & Fiedler, 2019)

According to the representation proposed here, deriflection would not be an issue any longer. Kinship for example, will become a gender on its own. In case of semantic agreement of the NCA, it will be considered as another NCM-cum-NCA with the possibility to specify the agreement of the NCA. Corbett (1991) considers that there is semantic 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" (Van de Velde, 2019). In other other words when the agreement is not semantic, it is syntactic. To understand the semantic agreement of the NCA, here is an example (Van de Velde, 2019, p.6):

Swahili G42 (Wald 1975: 241-242)

a. ki-le   ki-su,   ni-li-ki-on-a.
    PP7-DEM   7-knife   SP1SG-PST-OP7-see-FV
    'That knife, I saw it.'
b. Yu-le   ki-boko,   ni-li-mw-ona.
    PP1-DEM.   7-hippo   SP1SG-PST-OP1-see-FV
    'That hippo, I saw it.'

It is important to keep the table as simple as possible to avoid being too language specific and to be able to have a consistent approach of the diversity of Niger-Congo noun class systems, though it may increase the number of 'genders' ('typological cross-linguistic genders') because they gender/NCM-cum-NCA.

One table represents the general overview of the noun class system ("Main table"), the other one gives the detailed data ("To fill"). Every time an information is added in the main table, if necessary, it can be elaborated in the other table.

This table attempts to convey all the criteria and main features of a typological framework for Niger-Congo noun classes in the most readable way. In this respect, it is encompassing since every Niger-Congo noun class system can fit into it. It also allows flexibility since it gives a general view and gets more specific if needed. This helps keep the table readable and it can be used for multipurpose.

The following explanations present each column and aspect. They follow the order of columns in the main table:

- **Identification:** The identification number is only important for the table to be consistent when creating the typological cross-linguistic 'genders'. It consists of giving each row an arbitrary number before being able to give a number to the genders. The Identification number (ID) is preceded by the '#' symbol to emphasize its arbitrariness. These numbers can go up very high because each entry of each language will need a different Identification number.
- **Gender / NCM-cum-NCA:** In the first stage, when filling this table, the gender/NCM-cum-NCA will remain blank. This column can be filled once the analysis of enough data reveals typological numbers for Niger-Congo cross-linguistic genders/NCM-cum-NCA.
- **Affixation:** For the affixation column, it is possible to choose between « prefix », « suffix », « infix » and « circumfix » from a drop-down list.
- **NCM:** The NCM column is divided in two parts, the second one depending on the first one. The first column requires the user to choose between "Pairing" or "Single class". Selecting "Pairing" will create a SG and PL section

in the second column. While selecting "Single class" will not change anything. Once this has been done, the NCM needs to be filled in the table entitled "To fill".

- **_Allomorphy:_** In case of allomorphy for the NCM, it is possible to specify the number of allomorphs and choose the conditioning of the allomorphy by choosing between "Morphophonological rules", "General phonological rules" and "No allomorphy". It is then possible to fill the forms of the allomorphs and give the morphophonological rules in the table "To fill".
- **_Definiteness:_** Some noun class systems express the definiteness which impacts the form of the NCM. If it is the case, we can select "Yes", otherwise, "No". We can then fill (or not) the Definite NCM in the table "To fill". It will then appear in the "Main Table".
- **_Alliterative:_** The alliterative column is like the NCM column in the sense that it is divided in two parts and involves the table "To fill". The first part is a choice between "Yes" if it is an alliterative system and "No" if it is not. If "yes" is selected, it will automatically copy paste the NCM cell of this same line. On the other hand, if it is not an alliterative system, we will be redirected to the table "To fill". In which we will have to enter manually the NCA forms for each concord.
- **_NCA agreement:_** This column aims at showing the cases of Semantic agreement of the NCA. As seen previously, some NCM-cum-NCA will consist of syntactic agreement, for most cases, but sometimes, the agreement can be semantic. It is possible to specify the agreement of the NCA by choosing between "Semantic" or "Syntactic".
- **_Assignment:_** The assignment column consists in choosing between "semantic", "phonological" or "morphological" or "mixed". If it is a "Semantic" assignment, the "meaning" column must be filled. Same applies for the "Phonological" and "Morphological" assignments, that entails to fill the columns "Phonological rule" and "Morphological rule" respectively. In a perfect world, it would be with pre-made tags. If it is "Mixed", it is possible to fill the three of the columns in the table "To fill". The name of each filled columned will then appear in the "Main Table".
- **_Derivational use:_** The derivational use column consists in choosing between "Yes only", "Yes, not only", and "No". "Yes only" means this gender has no inherent members and only has a derivational use. "Yes, not only" means that the gender has inherent members but can also be used for derivational purpose. "No" means the gender only has inherent members thus has no derivational use.
- **_Hierarchy:_** The Hierarchy column refers to the possibility to add a noun class to another noun class. In this column, it is possible to choose between "Primary" which means that this is a primary noun class, or "Secondary" if it is a secondary noun class. Secondary noun classes can be affixed to Primary

noun classes, but the opposite is not possible. If the noun class system does not display a hierarchy, all noun classes will be considered "Primary". Hierarchy is not a main feature but can be interesting for some languages or research questions.

To make sure the table is clear, I used a color code. Yellow indicates the title of each column. The cells that need to be filled manually are light blue. The cells that cannot be filled, because the feature does not exist in the noun class system, can be turned black. For example, some Niger-Congo languages do not mark the noun class on the Verbal Phrase, the cells in the VP columns can be turned black. This color code helps the table remaining readable.

2.3. Purposes and Advantages

This database is multipurpose ; it can be used in many ways and it provides answers to a large range of questions. It allows to focus on different parts of a noun class system and to identify cross-linguistic differences and similarities.

The two tables in one allow to go from very general to very detailed which makes it readable and flexible because we can adapt it to a diversity of questions cross- and intra-linguistic.

An important goal of this framework is to be used as much for language description as for typological analysis.

2.3.1. Reporting

In the language documentation field, this framework is a clear, flexible and encompassing tool that allows to consider all the specificities of a noun class system to report on it. As such, it could be used by all linguists who study Niger Congo languages. The possibility to add as many columns as necessary (but not to delete any) gives the researcher a lot of freedom and allows to adapt to any specificities of a language.

2.3.2. Analysis

This framework offers a way to report but also to analyze Niger-Congo noun class systems in a typological approach. By allowing linguists to develop statistics, it enables us to find statistical and implicational intra-genetic universals. This gives a

predictive dimension to this framework. Using R studio, a statistical tool, to process the data and answer typological research questions is a possibility.

2.4. Limits

Like any other man-made system, this framework has limitations that we will discuss in this last section. The first limit that we anticipate involves the mandatory blank for genders/NCM-cum-NCA. The second limit is due to the use of Excel.

As said previously, this database first aims at revealing typological cross-linguistic 'genders' for Niger-Congo noun classes. The initial period might be messy due to the high number of entries to process. Indeed, each entry will need a different identification number. Thus, in the beginning, we will have to assign arbitrary numbers to each line before considering it a typological gender. Processing all the data in order to create non-arbitrary typological cross-linguistic genders/NCM-cum-NCA can be laborious. A possible and frustrating limit resides in the fact that no typological cross-linguistic genders/NCM-cum-NCA may be found in the end. However, the framework and the database can still have many other usages.

Excel's flexibility is both its best asset and its worst flaw. Indeed, the more open a software is, as Excel is, the more likely it is that a user will make changes that will compromise the integrity of the reporting that the tool provides. On the other hand, the more flexible the tool is, the more accessible it is and the more in-depth the analysis can be. To be robust, it would have to be fully compliant with the rules and standards, but this is almost impossible since the table is left up to each user who has the possibility of adding columns to consider the reality of each language.

The other limit involved by Excel (without macro) can be considered as a typological issue. Regarding the gender assignment, for the meaning as much as for the phonological and syntactic rules, it should be possible to select as many tags as necessary to convey the meaning, or phonological or morphological rules of the gender. Excel does not allow the use of mutually non-exclusive tags, only a drop-down list from which we can only choose one option. Free cells can be totally discretionary, left to the initiative of each user, while the options of a drop-down list are necessarily predefined. The use of completely free cells creates a problem of homogeneity, as certain similarities can be masked by different denominations, while the choices imposed by drop-down menus can be too restrictive to reflect the reality of any language. Nevertheless, it could be possible to homogenize these tags to process the data, but it will demand to restructure then recompile the data filled in the assignment column for each table. The process will have to take longer than if the use of tags was possible.

To sum-up these limits, most of them are related to the use of Excel or require extra workload. One solution, which needs further investigations, may be to use a relational database management system, which seems more suitable for this project. This type of software would guarantee a greater robustness, analyses not being likely to be affected by uncontrolled modifications, for example of columns. In a more sophisticated program, tables can be obsolete to use a multi-dimensional database. It supposes, however, more experience in computer coding to develop and maintain the framework than the use of Excel.

Conclusion

The fourth and last chapter of this paper is devoted to the establishment of the mandatory criteria and main features for a typological framework for Niger-Congo noun classes, and the proposition of a possible framework attempting to overcome the challenges and meet the criteria stated previously in this paper. Indeed, as we have seen previously, the existing levels of description of the language families vary a lot because of history, some language families having been studied more than others, and because of local contexts, some languages being widespread and others not. This situation leads to a profusion of definitions, to apparent oppositions which can be hardly reconciled and to a lack of a global vision of noun classes.

This typological framework should meet three criteria: it should be generic to encompass every Niger-Congo noun class system and noun class, it should be flexible to adapt to evolving data and it should be predictive to allow the use of data in a statistical form. To develop a typological approach, it seemed that the use of a database gathering the essential criteria of noun classes of all Niger-Congo languages would be a good way to bring together all available data in a single format that would solve most of the challenges we identified so far.

In this paper, we have developed a database prototype, using an Excel spreadsheet, divided in two tables. One table represents the general overview of the system, the other one gives detailed data. The same fields in the database must be filled in for all languages. This database can be used in many ways and can provide answers to a large range of research questions. This framework allows to focus on different parts of a noun class system and to identify cross-linguistic differences and similarities. It is also a good reporting tool that allows to consider all the specificities of a noun class system to report on it.

As such, it could be used as a basis by all linguists who study Niger-Congo noun classes and to share information structured in the same way. The possibility to add

as many columns as needed gives researchers a lot of freedom and allows them to adapt the frame to the specificities of any language while still mentioning the main features for Niger-Congo noun class systems.

Nevertheless, this framework has limits. On the one hand, at the end of the process, there is no assurance of obtaining enough representative and diverse data to completely avoid current biases and reveal typological cross-linguistic 'gender' for Niger-Congo languages. On the other hand, most of the limits we anticipate are linked to the software. Indeed, the more open a software is, the more likely it is that some user will make changes that will compromise the integrity of data and the analysis. The other limit involved by Excel can be considered as a typological issue. The use of completely free cells creates a problem of homogeneity, as certain similarities may be masked by different denominations, while the choices imposed by drop-down lists can be too restrictive to reflect the reality of all noun class systems.

One solution, which needs deeper investigation, may be to use a relational database management system instead of Excel. This type of software would still be generic, flexible, and predictive but it would guarantee a greater robustness, analyses not being likely to be affected by uncontrolled changes. However, this would require more experience in computer development to develop and maintain the framework than the use of Excel.

Conclusion

The methodology implemented to answer our Research Question: *What are the criteria and challenges to create a typogical framework for Niger-Congo noun classes?* allows us to gather a set of findings concerning noun classes in Niger-Congo languages.

Chapters 1 and 2 highlight the background of this research question, showing the necessity for linguists, as much descriptivists as typologists, to create a typological framework for Niger-Congo noun classes.

Not all language families have benefited from the same analyses, with the same depth, at the same time, hence the presence of a bias that can lead to the false idea that noun classes are more developed in certain families (typically in Bantu languages), simply because they have been more and earlier documented.

It is difficult to find in literature a simple and comprehensive definition of noun classes in Niger-Congo languages. Researchers use their own definitions or do not define noun classes at all.

Still, some features can be defined. Nouns are assigned to classes according to their referent, such as gender or form, often based on purely conventional features. They are grouped into different classes, marked by prefixes, suffixes or both. All members of a given class share the same affix.

The number of noun classes vary from language to language. Noun classes are usually accompanied by a system of agreement. In the case of an alliterative agreement system, the marker that affects the noun is repeated with other elements such as determiners, adjectives, or numbers. Including, in some cases, the verb. In a non-alliterative system, the form of agreement may be different from the affix that affects the noun.

Existing frameworks show that they are specific to each language so that it is difficult to draw common features. For instance, in the case of Bantu languages, based on the proto-Bantu system of noun classes, the classification ranges from 1 to 20, sometimes adding letters, while, in the case of Atlantic languages, noun class systems generally use morphemes, in the case of Gur languages, there is no consensus on the classification of noun classes and, in the case of Kru languages, their noun class system displays a very specific and different structure.

Three types of challenges emerge from these findings. The first challenge concerns the instability of the descriptions of noun classes like, for instance, Wolof which recently "gained" two noun classes in just one year. The second challenge is related

to the gathering in the same framework of all the specificities of Niger-Congo languages which may seem particularly distant. Finally, the third challenge is to find a way to match genders between languages, which may be based on semantics as well as phonology or morphology.

It appears that technology can offer a smart way to overcome these challenges and to develop a synchronic typological approach. A database gathering the criteria of noun classes of all Niger Congo languages would be a good way to bring together data in a single format. This framework should meet three criteria: it should be generic to encompass all Niger-Congo noun class systems and specificities, it should be flexible to adapt to evolutions and it should be predictive to allow statistics.

A prototype database, using an Excel spreadsheet, was developed. This table allows to get a big picture as well as detailed information. The same fields being filled in for all languages, it can provide answers to a large range of research questions. This framework allows to focus on different parts of a noun class system and to identify cross-linguistic differences and similarities. It could be used by all linguists to gather information on Niger-Congo noun classes and to share information structured in the same way. The possibility to add as many columns as needed allows researchers to adapt to the specificities of any noun class system.

However, we can imagine that the use of such a system will have its limits. To be representative, it assumes that a significant amount of data, from all Niger-Congo languages, is collected, which is not assured. In addition, Excel technology has its own limitations because of its flexibility, which does not guarantee against the risks of data corruption and erroneous processing. A workaround could be to use a real database management system instead of Excel. The solution we outline, based on technology, is at this stage only a simple proposal. It is a first step towards the constitution of a typological framework for Niger-Congo noun classes that will have to be tested and enriched by the linguist community. This prototypic typological framework for Niger-Congo noun class systems takes into account the criteria and attempts to overcome the challenges stated in this paper as much as possible.

References

Alexandre, P. (1959). Développements récents des études bantu à Londres. *Journal des Africanistes*, *29*(2), 297–304.

Babou, C. A., & Loporcaro, M. (2016). Noun classes and grammatical gender in Wolof. *Journal of African Languages and Linguistics*, *37*(1), 1–57. https://doi.org/10.1515/jall-2016-0001

Barlow, M., & Ferguson, C. A. (1988). *Agreement in natural language: Approaches, theories, descriptions*. Center for the Study of Language and Information.

Barnwell, K. G. L. (1969). *A grammatical description of Mbembe (adun dialect): A Cross River language*. Department of General Linguistics, University College London.

Bendor-Samuel, J. (2006). Niger-Congo Languages. In K. Brown (Ed.), *Encyclopedia of Language & Linguistics (Second Edition)* (pp. 631–635). Elsevier. https://doi.org/10.1016/B0-08-044854-2/02301-4

Bendor-Samuel, J. (2012). *Kordofanian languages*. Encyclopedia Britannica. https://www.britannica.com/topic/Kordofanian-languages

Bendor-Samuel, J. (2020). *Gur languages*. Encyclopædia Britannica Inc. http://academic.eb.com/levels/collegiate/article/75705

Bendor-Samuel, J., & Hartell, R. L. (1989). *The Niger-Congo languages: A classification and description of Africa's largest language family*. University Press of America.

Bleek, W. H. I. (1862). *A Comparative Grammar of South African Languages*. London; Cape Town [printed].

Blevins, J. (2013). The information-theoretic turn. In *Psihologija* (pp. 355–375).

Bobuafor, M. (2013). *A grammar of Tafi*. LOT.

Bole-Richard, R. (1983). *Systématique phonologique et grammaticale d'un parler Ewe: Le Gen-Mina du Sud-Togo et Sud-Bénin*. Editions L'Harmattan.

Boutwell, R. L. (2020). *A grammar of Nchane: A Bantoid (Beboid) language of Cameroon* [PhD Thesis].

Carlson, R. (1994). *A grammar of Supyire*. Mouton de Gruyter.

Childs, G. T. (2011). A Grammar of Kisi: A Southern Atlantic Language. In *A Grammar of Kisi*. De Gruyter Mouton. http://www.degruyter.com/view/title/3852

Cole, D. T. (1959). *Doke's classification of Bantu languages*.

Contini-Morava, E. (2002). (What) do noun class markers mean? In W. Reid, R. Otheguy, & N. Stern (Eds.), *Studies in Functional and Structural Linguistics* (Vol. 48, pp. 3–64). John Benjamins Publishing Company. https://doi.org/10.1075/sfsl.48.04con

Corbett, G. G. (1991). *Gender*. Cambridge University Press. https://doi.org/10.1017/CBO9781139166119

Corbett, G. G. (2006). *Agreement*. Cambridge University Press.

Corbett, G. G., & Fraser, N. M. (2000). Gender assignment: A typology and a model. In *Systems of Nominal Classification* (Gunter Senft, pp. 293–325). Cambridge University Press.

Creissels, D. (1991). *Description des langues négro-africaines et théorie syntaxique*. ELLUG, Université Stendhal.

Creissels, D. (1999). "Genres" indo-européens et "classes nominales" Niger-Congo. *Faits de langues*, *7*(14), 177–184. https://doi.org/10.3406/flang.1999.1280

Creissels, D. (2001). Les systèmes de classes nominales des langues Niger-Congo: Prototype et variations. *Linx. Revue des linguistes de l'université Paris X Nanterre*, *45*, 157–166. https://doi.org/10.4000/linx.837

Creissels, D. (2014). Derivation in Niger-Congo. In *The Oxford Handbook of Derivational Morphology* (Rochelle Lieber and Pavol Stekaur). https://www.academia.edu/14548515/Derivation_in_Niger_Congo

Creissels, D. (2015). *Noun class systems in Atlantic languages*. 31.

Creissels, D., & Pozdnjakov, K. I. (2015). *Les classes nominales dans les langues atlantiques*. Rüdiger Köppe Verlag.

Daull, A. (1879). *Grammaire de kisouahili*.

de Wolf, P. (1971). The Noun-Class System of Proto-Benue-Congo. In *The Noun-Class System of Proto-Benue-Congo*. De Gruyter Mouton. https://www-degruyter-com.ezproxy.leidenuniv.nl/document/doi/10.1515/9783110905311/html

Devos, M. (2008). *A grammar of Makwe: (Palma; Mozambique)*. Lincom Europa.

Dixon, R. M. W. (1986). Noun Classes and Noun Classification in Typological Perspective. In *Noun Classes and Categorization: Proceedings of a Symposium on Categorization and Noun Classification* (pp. 105–112). http://web.b.ebscohost.com.ezproxy.leidenuniv.nl:2048/ehost/ebookviewer/ebook/ZTAwMHh3d19fMzk1MzM5X19BTg2?sid=d3a7a43c-f117-4400-9934-6c5eb640b97b@pdc-v-sessmgr02&vid=0&lpid=lp_105&format=EB

Dorvlo, K. (2008). *A grammar of Logba (Ikpana)*. LOT.

Dube, P., Ndebele, L., & Ndlovu, M. (2014). An analysis of the status of the secondary noun prefixes in Ndebele. *South African Journal of African Languages*, *34*(2), 145–149. https://doi.org/10.1080/02572117.2014.997050

Gibbard, G., Rohde, H., & Rose, S. (2009). *Moro Noun Class Morphology*. 12.

Godé, V. (2008). *Le dadjriwalé: Langue kru de la Côte d'Ivoire*. L'Harmattan.

Good, J., & Creissels, D. (2018). Current issues in African morphosyntax. In *The Languages and Linguistics of Africa* (pp. 709–881). De Gruyter Mouton. https://doi.org/10.1515/9783110421668-007

Greenberg, J. H. (1955). *Studies in African linguistic classification*. Compass.

Greenberg, J. H. (1963). Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In *Universals of Human Language* (Greenberg, Joseph H., pp. 73–113).

Grinevald, C. (1999). Typologie des systèmes de classification nominale. *Faits de langues*, *7*(14), 101–122. https://doi.org/10.3406/flang.1999.1271

Guest, E. (1997). *Heiban Noun Classes*.

Güldemann, T., & Fiedler, I. (2019). Niger-Congo "noun classes'" conflate gender with deriflection. In *Grammatical gender and linguistic complexity, volume 1: General issues and specific studies* (pp. 95–145). Zenodo. https://doi.org/10.5281/ZENODO.3462762

Gutman, A., & Avanzati, B. (2013). *Niger-Congo Languages*. The Language Gulper. http://languagesgulper.com/eng/Niger.html

Harley, M. (2005). *A grammar of Tuwuli*. https://www.academia.edu/15809233/A_grammar_of_Tuwuli

Heine, B. (2006). African Linguistics: History. In K. Brown (Ed.), *Encyclopedia of Language & Linguistics (Second Edition)* (pp. 94–95). Elsevier. https://doi.org/10.1016/B0-08-044854-2/04377-7

Hepburn-Gray, R. (2016). *A survey of Niger-Congo noun class agreement systems*. 3.

Hepburn-Gray, R. (2017). *Against split morphology: 'Inflection', 'derivation', and the structure of the SBCG lexicon*. 24th International Conference on Head-driven Phrase Structure Grammar. https://doi.org/10.1007/978-94-017-3712-8_2

Hepburn-Gray, R. (2018). *Niger-Congo Noun Classes and the Typology of Gender*.

Hinnebusch, T. (1989). Bantu. In *The Niger-Congo Languages* (Bendor-Samuel, pp. 450–473).

Idiata, D. F. (1998). *Aperçu sur la morphosyntaxe de la langue isangu (Bantou, B42)*. LINCOM Europa.

Joseph, B. Y. (2009). *La Structure de DP dans les langues Kwa*. 28.

Kamarah, U. I. (2007). *A descriptive grammar of KaThemne (Temne)*. Lincom Europa.

Kantchoa, L. (2014). *La classification nominale du gidere, langue volta-mono du Togo et du Ghana*. *00*, 13.

Kasonde, A. R. M. (2009). *Phonologie et morphologie de la langue bemba*. LINCOM Europa.

Kießling, R. (2013). On the origin of Niger-Congo nominal classification. In R. Kikusawa & L. A. Reid (Eds.), *Current Issues in Linguistic Theory* (Vol. 326, pp. 43–65). John Benjamins Publishing Company. https://doi.org/10.1075/cilt.326.05kie

Kipré, F. B. (2016). Construction directe » vs « construction indirecte »:quelle interprétation sémantique dans la relation de possession en bété ?

Cognition, Représentation, Langage, *14*(14‑2).
https://doi.org/10.4000/corela.4631

Koelle, S. W. (1963). *Polyglotta africana* (Unveränd. Nachdruck / verm. durch eine historische Einführung von P.E.H. Hair und einen Wortindex von David Dalby.). Akademische Druck- uVerlagsanstalt.

Konoshenko, M., & Shavarina, D. (2019). A microtypological survey of noun classes in Kwa. *Journal of African Languages and Linguistics*, *40*(1), 75–114. https://doi.org/10.1515/jall-2019-0004

Kra, K. A. E. author. (2016). *Le koulango, langue gur de Côte d'Ivoire et du Ghana*. L'Harmattan.

Krasnoukhova, O. (2019). Linguistic Typology. In *The SAGE Encyclopedia of Human Communication Sciences and Disorders* (Vol. 1–4, pp. 1106–1108). SAGE Publications, Inc. https://doi.org/10.4135/9781483380810

Laughlin, F. M. (1997). *Noun Classification in Wolof: When affixes are not renewed*. 28.

Levi, J. A. (2009). Portuguese and Other European Missionaries in Africa A look at their linguistic production and attitudes (14151885). *Historiographia Linguistica*, *36*(2), 363–392. https://doi.org/10.1075/hl.36.2.10lev

Lumwamu, F. (1970). Sur les classes nominales et le nombre dans une langue bantu. *Cahiers d'études africaines*, *10*(40), 489–529. https://doi.org/10.3406/cea.1970.2818

Maho, J. F. (1999). *A comparative study of Bantu noun classes*. Acta Universitatis Gothoburgensis.

Marchese, L. (1988). Noun Classes and Agreement Systems in Kru: A Historical Approach. In *Agreement in natural language: Approaches, theories, descriptions*.

Marchese Zogbo, L. (2005). *Kru revisited, Kru revealed*.

McCormack, A. (2007). *Subject and Object Pronominal Agreement in the Southern Bantu languages: From a Dynamic Syntax Perspective*. https://www.cambridge.org/core/product/identifier/S0001972000021033/type/journal_article

Meeussen, A. E. (1967). Bantu grammatical reconstructions. *Africana Linguistica*, *3*(1), 79–121. https://doi.org/10.3406/aflin.1967.873

Meeuwis, M. (2010). *A grammatical overview of Lingála*. LINCOM EUROPA.

Meinhof, C. (1932). *Introduction to the phonology of the Bantu languages*.

Miehe, G., & Winkelmann, K. (2007). *Southwestern Gur languages (without Gurunsi)*. Köppe.

Morrison, M. E. (2018). Beyond derivation: Creative use of noun class prefixation for both semantic and reference tracking purposes. *Journal of Pragmatics*, *123*, 38–56. https://doi.org/10.1016/j.pragma.2017.10.009

Mouguiama-Daouda, P. (2005). *Contribution de la Linguistique à L'histoire des Peuples du Gabon La Méthode Comparative et Son Application Au Bantu*. Editions du Centre national de la recherche scientifique.

Msaka, P. K. (2019). *Nominal classification in Bantu revisited: The perspective from Chichewa*. 359.

Musah, A. A. (2018). *A grammar of Kusaal: A Mabia (Gur) language of northern Ghana*. Peter Lang.

Nambima, Y. T. (2017). *Sémantique des classes nominales du̅ m̄-bɛ̄li-mɛ*. Université d'Abomey-Calavy.

N'dré, D. (2016). Les mécanismes de construction de mots en godié, langue kru de Côte D'Ivoire. *Linguistik Online*, *77*(3), Article 3. https://doi.org/10.13092/lo.77.2905

Nicole, J. (1999). *Les classes nominales dans les langues voltaïques: Esquisse d'un cadre de description*. SIL Togo.

Obeng, S. (2008). *Efutu grammar*. Lincom Europa.

Okhotina, N. V. (1975). *Some Structural Changes in Bantu Languages Due to Their Specific Communicative Functions* [Working Paper]. Ohio State University. Department of Linguistics. https://kb.osu.edu/handle/1811/81831

Platiel, S., & Kaboré, R. (1998). *Les langues d'Afrique subsaharienne*. Ophrys.

Pozdniakov, K., & Robert, S. (2015). *Les classes nominales en wolof: Fonctionnalités et singularités d'un système restreint*. ResearchGate. https://www.researchgate.net/publication/297777429_Les_classes_nominales_en_wolof_fonctionnalites_et_singularites_d'un_systeme_restreint

Rickard, D. T. (1970). *Kru grammar: (A preliminary study)*. Department of Literacy Work and Literature Production of the United Methodist Church.

Sambiéni, C. (2014). Sémantique multidimensionnelle et fonctions des classes nominales: L'exemple du biali, langue gur Oti-Volta-orientale, Bénin. *Journal of African Languages and Linguistics*, *35*(2), 205–241. https://doi.org/10.1515/jall-2014-0007

Sapir, E. (1907). *Notes on the Takelma Indians of Southwestern Oregon*.

Sapir, J. D. (1965). *A grammar of Diola-Fogny: A language spoken in the Basse-Casamance region of Senegal*.

Segerer, G. (2015). *Les classes nominales en keerak (joola de Kabrousse)*. 47.

Semantique_des_classes_nominales_du_m_b.pdf. (n.d.).

Smits, H. (2017). *A grammar of Lumun. A Kordofanian language of Sudan. Vol. I.* LOT, Netherlands Graduate School.

Soukka, M. (2000). *A descriptive grammar of Noon: A Cangin language of Senegal*. LINCOM EURPOA.

Steele, S. (1978). Word order variation: A typological study. In *Universals of Human Language* (Joseph Harold Greenberg, Charles Arlbert Ferguson&Edith A. Moravcsik, pp. 585–623). Stanford University Press.

Traoré, Y., & Féry, C. (2018). *Nominal classes and phonological agreement in Tagbana*. https://escholarship.org/uc/item/7h94z08p

Van de Velde, M. (2019). Nominal morphology and syntax. In M. Van de Velde, K. Bostoen, D. Nurse, & G. Philippson (Eds.), *The Bantu Languages* (2nd ed., pp. 237–269). Routledge. https://doi.org/10.4324/9781315755946-8

van den Eynde, K., & Mufuta, P. (1994). Le système de référence primaire du ciluba (luba-Kasai): Analyse en traits. *International Journal of Lexicography*, *7*(2), 90–105. https://doi.org/10.1093/ijl/7.2.90

Vanderelst, J. (2016). *A grammar of Dagik: A Kordofanian language of Sudan*. Rüdiger Köppe Verlag.

Voll, R. M. (2017). *A grammar of Mundabli: A Bantoid (Yemne-Kimbi) language of Cameroon* [Doctoral Thesis, Leiden University]. https://openaccess.leidenuniv.nl/handle/1887/56258

Welmers, W. E. (1973). *African language structures*. University of California Press.

Williamson, K. (1989). Niger-Congo Overview. In *The Niger-Congo languages* (John Bendor-Samuel, pp. 3–45).

Williamson, K., & Blench, R. (2000). Niger-Congo. In *African Languages: An Introduction* (Bernd Heine&Derek Nurse, pp. 11–42).

Wit, G. de. (2015). *Liko phonology and grammar: A Bantu language of the Democratic Republic of the Congo* [PhD Thesis].

Zogbo, R. G. (2004). *Parlons bété: Une langue de Côte d'Ivoire*. L'Harmattan.