VERBAL CATEGORIES IN NIGER-CONGO LANGUAGES

Derek Nurse, Sarah Rose and John Hewson

(with collaboration from Christa Beaudoin-Lietz)

List of Contents

Introduction

0.1 Purpose
0.2 Niger-Congo, the Niger-Congo languages, the choice of languages
    0.2.1 Language sample
0.3 The structure of the book and of the chapters
0.4 Morphosyntactic background
    0.4.1 Word order in Niger-Congo
    0.4.2 The structure(s) of verbs in Niger-Congo
    0.4.3 Serial verbs
    References for Introduction

Chapter 1 Aspect and Tense

1.1 Aspect
    1.1.1 Perfective, Perfect and Factative
    1.1.2 Incompletives
1.2 Tense
    1.2.1 Aspect and tense
    1.2.2 Aspect systems and "future tense(s)"
    1.2.3 Tense systems
1.3 A brief note on terminology ~ Sarah Rose
    References for Chapter 1

Chapter 2 Aghem

2.1 General
2.2 Word Order
2.3 Verb structure
2.4 Tense, aspect
2.5 Other categories
2.6 Negation
2.7 'Be' and 'have', and sources for other formatives
2.8 Bamileke-Dschang
    References for Chapter 2

Chapter 3 Bambara

3.1 General
3.2 Word Order
3.3 Verb structure
3.4 Auxiliaries
3.5 Participles
3.6 Prefixes and suffixes
3.7 Mood
3.8 Negatives

3.9 Relatives
3.10 Conclusion
References for Chapter 3

Chapter 4 Bantu Narrow

4.1 General
4.2 Word Order
4.3 Verb structure
4.4 Tense, aspect
4.5 Other categories: relative, mood, focus
4.6 Negation
4.7 Auxiliaries
References for Chapter 4

Chapter 5 Bijago

5.1 General
5.2 Word Order
5.3 Verb structure
5.4 Multiple verb constructions
5.5 Negatives
References for Chapter 5

Chapter 6 Degema

6.1 General
6.2 Word order
6.3 Verb structure
6.4 Aspect
6.5 Mood
6.6 Auxiliaries
6.7 Negation
References for Chapter 6

Chapter 7 Donnɔ Sɔ (Dogon)

7.1 General
7.2 Word order
7.3 Verb structure
7.4 Participles
7.5 Aspect, Tense and Mood
7.6 Verbal forms
7.7 Auxiliaries
7.8 Negation
7.9 Summary of forms
References for Chapter 7

Chapter 8 Doyayo

8.1 General
8.2 Word order
8.3 Word Formation
8.4 Verb structure
8.5 The Representation of Modality
8.6 Negatives
References for Chapter 8

Chapter 9 Ejagham

9.1 General
9.2 Word order
9.3 Verb structure
9.4 Aspect
9.5 Other Verbal Categories
9.6 Negatives
9.7 Auxiliaries ‘come’ and ‘go’
9.8 Focus
9.9 Conclusions
9.10 Diagrammatic Representations
References for Chapter 9

Chapter 10 Ewe

10.1 General
10.2 Word Order
10.3 Verb Structure
10.4 Aspects and a “tense/mood”
10.5 Other categories
10.6 Negation
10.7 ‘Be’, “co-verbs”, auxiliaries
References for Chapter 10

Chapter 11 Fula

11.1 General
11.2 Word order
11.3 Verb structure
11.4 Aspect, Mood, Voice, Negative (AMVN)
11.5 Simple forms (forms which do not contain an auxiliary)
11.6 Complex forms (forms which do contain an auxiliary)
11.7 Mood
11.8 Non-finite forms
11.9 Negation
11.10 Focus
11.11 Overview of Fula verb forms

References for Chapter 11

Chapter 12 Godié

12.1 General
12.2 Word Order
12.3 Verb structure
12.4 TAM categories
12.5 Other categories: Subjunctives
12.6 Negatives
12.7 Auxiliaries
References for Chapter 12

Chapter 13 Ijo

13.1 General
13.2 Word order
13.3 Verb structure
13.4 Tense, aspect
13.5 Other categories
13.6 Negation
13.7 'Be'-verbs, auxiliaries
References for Chapter 13

Chapter 14 Jukun

14.1 General
14.2 Word Order
14.3 Verb structure
14.4 Aspect, mood
14.5 Other categories
14.6 Negation
14.7 'Be', auxiliaries, and modals
References for Chapter 14

Chapter 15 Kabiye

15.1 General
15.2 Word Order
15.3 Verb Structure
15.4 Aspect
15.5 Auxiliaries
15.6 Other Categories
15.7 Negatives
15.8 Relatives
References for Chapter 15

Chapter 16 Kisi

16.1 General
16.2 Word order
16.3 Verb structure
16.4 Aspect, tense?, mood
16.5 Auxiliaries
16.6 Relatives
16.7 Negation
16.8 Overview of Kisi verb forms
References for Chapter 16

Chapter 17 Makaa

17.1 General
17.2 Word Order
17.3 Verb Structure
17.4 Aspect, Mood
17.5 Other Categories
17.6 Negation
17.7 Diagram of tense-aspect system
References for Chapter 17

Chapter 18 Obolo

18.1 General
18.2 Word Order
18.3 Verb structure
18.4 Aspect, mood
18.5 Other categories
18.6 Negation
18.7 ‘Be’, auxiliaries
References for Chapter 18

Chapter 19 Otoro

19.1 General
19.2 Word order
19.3 Verb structure
19.4 Compound verbs: locative ‘be’ and main verb = Progressive
19.5 Pronominal subject and object marking
19.6 Aspect
19.7 Other verbal categories
19.8 Summary of co-occurrence of (suffixal) FAC, IPFV/VEN, SBJ with other categories
19.9 Negation
References for Chapter 19

Chapter 20 Supyire

20.1 General
20.2 Word Order
20.3 Verb structure
20.4 Tense, aspect, mood
20.5 Serial verbs
20.6 Other categories
20.7 Negation
20.8 Auxiliaries
References for Chapter 20

Chapter 21 Yoruba

21.1 General
21.2 Word Order
21.3 Verbal Structure
21.4 Tensed verbs with Performative, Progressive, and Perfect
21.5 Auxiliaries
21.6 Expressing past and present
21.7 Modal forms
21.8 Negatives
21.9 Conclusion
References for Chapter 21

Chapter 22 Zande

22.1 General
22.2 Word order
22.3 Verb description and structure
22.4 Pre-stem markers
22.5 Tense distinctions
22.6 Subjunctive and Imperative
22.7 Negation/Relativization
22.8 Conclusion
References for Chapter 22

List of Abbreviations

1, 2, 3	First, second, and third person, respectively. Also noun class numbering
A	aspect
A, B, C….	Guthrie's 15 zones: A, B, C, D, E, F, G, H, K, L, M, N, P, R, S. A listing such as A, (B), C means 'attested in all A and C languages but only some B languages'.
ADV	adverb or adverbial
AM	aspect, mood
APP	applicative
ASP	aspect

AUX		auxiliary
BEN		beneficiary
C		consonant
CARP		acronym for the commonest (neutral?) ordering of the four commonest extensions (CAU, APP, REC, PAS). Devised by L. Hyman.
CAU		causative
CFL		counterfactual
Cl		class(es) or class marker
CM		clause marker
CND		conditional
CNJ		conjunctive
CON		conditional
CTV		connective
COP		copula
DEF		definite
DEIC		deictic
DIR		directional
DO		direct object
DUR		durative
EMPH		emphatic
ENC		enclitic
EXCL		exclamatory (particle), exclamation
EXT		extension
F		final (includes final vowels, CV suffixes, etc)
FAC		factative
FOC		focus or focus marker
FUT		future
	IMM	immediate (future)
	HOD	hodiernal (future)
	MID	middle (future)
	FAR	far (future)
	F_1, F_2, F_3, F_4	degrees of future distance from the present, F_1 being the closest, F_4 the Farthest
FP		future prefix
FV		final vowel
GEN-N		genitive followed by a nominal
H		high tone(d) (or zone H, as above)
HAB		habitual
HYP		hypothetical
IMP		imperative
IND		indicative
INCE		inceptive
INCH		inchoative
INF		infinitive
INTRANS		intransitive
INST		instrumental

IO	indirect object
IP	intransitive prefix
IPFV	imperfective
IRR	irrealis
ITR	iterative
ITV	itive
L	low tone(d) (or zone L, as above)
LOC	locative
M	mood
M	mid (tone)
MOD	modal, modality
N	nasal, realised as [m, n, ŋ, ɲ], depending on place of following segment
NAR	narrative
NC	Niger-Congo
NEG	the category negative, or the position in the word, or negation
NF	near future
NP	noun phrase
O/OBJ	object
OM	the pre-stem Object Marker; object marker (affixed)
OP	object pronoun (independent of verb)
OPT	optative
P	plural (1p = first person plural, etc)
P	past (tense)
PART	participle
PAS	passive
PAST	
IMM	immediate (past)
HOD	hodiernal (past)
HES	hesternal (past)
MID	middle (past)
FP	far = remote past
P_1, P_2, P_3, P_4	degrees of past distance from present, P_1 being the closest, P_4 the farthest
PER	persistive
PFM	performative
PFT	perfect
PFV	perfective
POT	potential
Pre-SM or PreSM	the position before SM in the verbal string
PRG	progressive
PRH	prohibitive
PRO	pronoun
PROS	prospective
PRS	present
RCP	reciprocal
REL	relative, pronoun or marker
REM	remote

REP	repetitive
RES	resultative
S	singular (1s = first person singular, 2s, 3s)
S	subject (or zone S, as above)
SBJ	subjunctive
SBS	subsecutive
SC	subject concord (= SM)
SIM	simultaneous
SIT	situative
SM	subject marker (= SC) (affixed)
SP	subject pronoun (independent of verb stem)
STAT	stative
SUF	suffix
SVO	subject verb object
T	tense
TA(M)	tense-aspect(-mood)
TM	tense marker
TRANS	transitive
V	verb
V	vowel
VB	verb, or verbal, or verbal base
VEN	ventive
VOL	volitional
V-NOM	verb plus nominalizer
X	any constituent occurring after the O or V as a sentence constituent, most often referring to adverbials

Other Abbreviations

AAP = Afrikanistische Arbeitspapiere
ACAL = Annual Conference on African Linguistics
AL = Africana Linguistica
AU = Afrika und Übersee
BLS = Berkeley Linguistics Society
(B)SOAS = Bulletin of the School of Oriental and African Studies
CNRS = Centre National de la Recherche Scientifique
CNWS = Center for non-Western Studies. Leiden
CUP = Cambridge University Press
EA = Estudias Africanos
IAI = International African Institute
JALL = Journal of African Languages and Linguistics
JWAL = Journal of West African Linguistics
LLACAN = (France)
MRAC = Musée Royal de L'Afrique Centrale
OGMIOS [Name of the Newsletter for the Foundation of Endangered Languages] www.ogmios.org/home

OUP = Oxford University Press
RKV = Rüdiger Köppe Verlag
SAL = Studies in African Linguistics
SCOPIL = Southern California Occasional Papers in Linguistics
SELAF = Société d'Etudes Linguistiques et Anthropologiques de France
SIL = Summer Institute of Linguistics (International)
SOAS = School of Oriental and African Studies
UCB = University of California at Berkeley
UCLA = University of California at Los Angeles
USC = University of Southern California
WAJL = West African Journal of Linguistics.
WOCAL = World Conference on African Linguistics

Conventions

=	clitic boundary
-	indicates affixation; morpheme boundary
	space between two items indicates they are separate words
#	word boundary
/.../	underlying or phonemic form
[...]	phonetic form
*	reconstructed or proto form (usually PB)
+	'and' or 'plus'
>	becomes, became
ꜜ	(tone) downstep

Introduction

0.1 Purpose

This book is conceived as a sequel to *Tense and Aspect in Bantu* (Nurse 2008). That book concentrated on the typology of tense and aspect (henceforth TA) across a wide and representative set of (Narrow) Bantu languages. It aimed to establish the range within which Bantu languages vary in their grammaticalized expression of TA, how tense and aspect interact, their semantic content, and to some extent their pragmatics. It also examined other verbal categories but in less depth.

Our purpose here is similar. Since we were familiar with verbal categories in Bantu, we were curious about the same categories in wider Niger-Congo and about the general Niger-Congo background from which Bantu had emerged. How far were those categories and that background similar? We discovered many similarities and we also found significant differences. For instance, whereas all Bantu languages encode both aspect and tense, a majority of Niger-Congo families can best be analyzed in terms of aspect alone. Bantu and wider Niger-Congo also differ structurally - whereas what precedes the verb stem in Bantu languages has a synthetic structure, many Niger-Congo languages have a string of discrete items before the stem, so their structures can be viewed as analytic.

Our purpose can be expressed more broadly. We saw that no one had attempted to present an overview of verb systems in Niger-Congo, and we are aiming to fill that gap. Our main focus is aspect (and tense) but we also document word order, verb structure, mood, imperatives, focus, relativization, negation, and auxiliary verbs, particularly 'be' verbs, because they are important sources of aspect in Niger-Congo.

While our main purpose is typology, that is, the system of verbal categories, their architecture and meanings, we also deal with description, diachrony and reconstruction, and change and grammaticalisation.

We present analyses of a set of languages that include a descriptive component because we have found that most of our correspondents and colleagues, Africanist and general, know little of verbs and verb categories in Niger-Congo in general. Niger-Congo is so huge that some concentration is necessary, so scholars are typically familiar with one language or the languages of one family within Niger-Congo but not with the facts of other families. We make a conscious effort to present material in a way that we hope will raise readers' awareness of what is to be found in Niger-Congo.

Our analysis further includes a component on diachrony and reconstruction because of our interest in comparing the facts of Bantu with those of wider Niger-Congo. From asking which systems, categories and structures occur today, we move to asking which might be assumed to be older. Inevitably, this raises the issue of change, because if it is possible to show that certain features are older, then we have to ask what has changed and why. Central to this is grammaticalisation: many of the features we assume are innovations, as verbal inflections or clitics, derived from the incorporation of auxiliary verbs and other independent items.

This book will be of interest to those interested in the theory and practice of verb categories and systems, to Africanists and Niger-Congo specialists, graduate and advanced students, and will be a useful a reference tool for libraries and individuals.

0.2 Niger-Congo, the Niger-Congo languages, the choice of languages

Niger-Congo is a huge language phylum. During the second half of the last century the 'lumping' view prevailed, which saw Niger-Congo as a genetic unit, still the mainstream opinion. This Niger-Congo consists of around 1500 languages (Gordon 2005), making it the largest phylum in the world and in Africa[1]. Roughly five hundred languages are Bantu, the remaining thousand non-Bantu. Some 'splitters' have recently voiced doubts about the membership of some families in Niger-Congo: Güldemann (2008) excises Atlantic, Dogon, Ijoid, and Mande, to which Dimmendaal (2008) adds Ubangi. If all five are excluded, then Niger-Congo loses its status as the world's largest phylum; if Ubangi is retained, then it remains the largest. Resolving classificatory disputes [of this kind] depends on considering all available linguistic evidence, which we do not do here. Based on verbal evidence, we incline to the mainstream view of Niger-Congo as a unit, and we reproduce below Blench's most recent (p.c.) diagrams of Niger-Congo (1a) and Benue-Congo (1b) (cf Williamson & Blench (2000:18).

It can be seen from this diagram alone that it is hard to say how many families make up Niger-Congo. Further, the subset at the bottom left of Diagram 1a (and then (1b), Benue-Congo, contains well over 900 languages, divided into various subsets, including eventually Narrow Bantu. What is family, what is sub-family, what is group or subset? This is relevant to our task of dealing with verbal categories in Niger-Congo. Nurse (2008) presented detailed data from a hundred Bantu languages and took data from another hundred or so, out of a total of some five hundred, so could be reasonably certain that the total (some 40%) sample represented a typological and geographical cross section of Bantu. In our case, how many of the one thousand or so non-Bantu Niger-Congo languages would constitute a reasonable typological and geographical sample? Many Niger-Congo languages, spoken by small and often dwindling communities, are not described or are underdescribed. This not only immediately limits our choice but also makes judgements about typicality difficult: if a family or group consists of many languages, of which only one is well described, while the others are not described, how are we to know if that one is typical and could represent the others well? Our solution was necessarily simple and arbitrary. We chose one language from each family as shown in Diagram 1, and then one language from groups within some of the larger 'families' such as Adamawa-Ubangi or Benue-Congo. The choice of representative language was made largely on the basis of the availability of a description or analysis, or in some cases more than one description or analysis of the language. In some cases, we were able to communicate electronically with authors. Since it is often, but not always, the languages of larger communities that have been described, we run the risk of presenting the verbal systems of languages that have been simplified by long use as lingua francas. At the same time, since they are used by large numbers of people, there is often considerable current dialect variation (e.g. in Fula), and in that case, which variant are we to describe? We recognize these limitations and we acknowledge that twenty-one languages may be inadequate representation of the thousand or so non-Bantu Niger-Congo languages.

[1] Over 400 million Africans, over half the continent's population, speak a Niger-Congo language.

Figure 1. **Niger-Congo restructured (Data in both figures from Blench)**

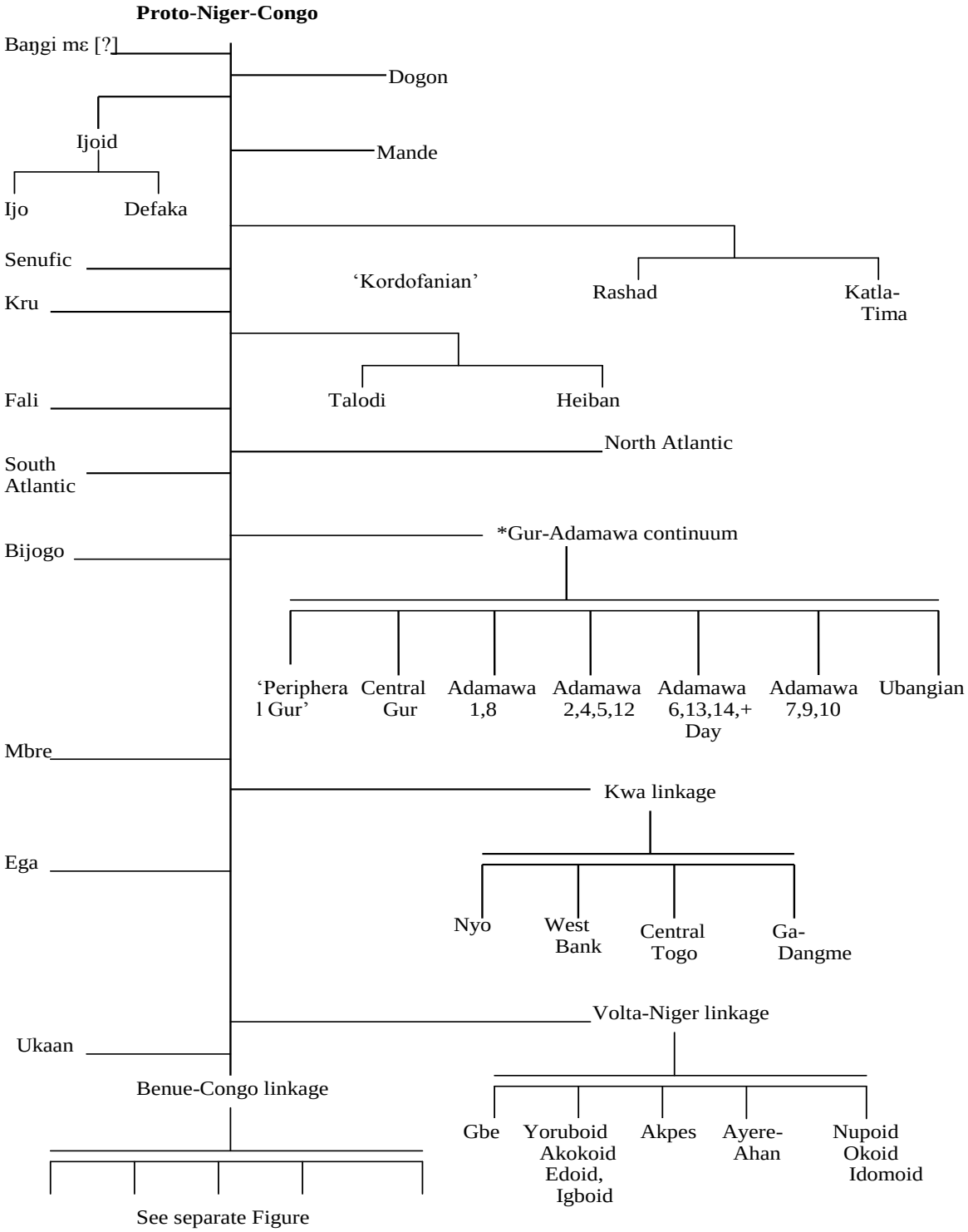
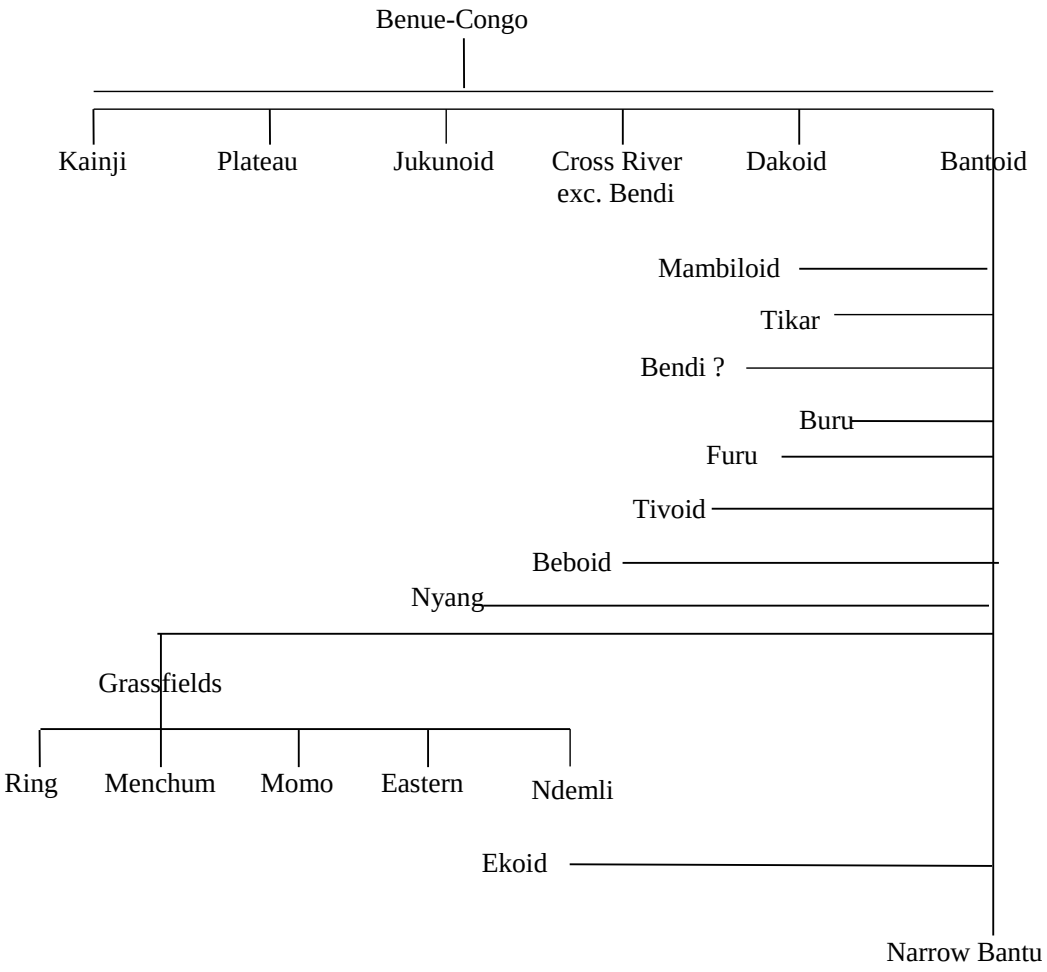


Figure 2. **Revised subclassification of Benue-Congo languages**



0.2.1 Language Sample

Our language sample is, language first, 'family' in brackets, followed by the Chapter which discusses each:

Aghem (Grassfields Bantu, Bantoid, Benue-Congo)	Chapter 2
Bambara (Mande)	Chapter 3
Bantu (general overview of Narrow Bantu, Bantoid, Benue-Congo)	Chapter 4
Bijago (formerly part of Atlantic)	Chapter 5
Degema (Edoid, Volta-Niger)	Chapter 6
Donno So (Dogon)	Chapter 7
Doyayo (Adamawa, Gur-Adamawa-Ubangi)	Chapter 8
Ejagham (Ekoid, Bantoid, Benue-Congo)	Chapter 9
Ewe (Kwa)	Chapter 10
Fula (North Atlantic)	Chapter 11
Godie (Kru)	Chapter 12
Ijo (Ijoid)	Chapter 13
Jukun (Benue-Congo)	Chapter 14
Kabiye (Gur, Gur-Adamawa-Ubangi)	Chapter 15
Kisi (South Atlantic)	Chapter 16
Makaa (northwest Narrow Bantu, as Ch 4)	Chapter 17
Obolo ((Lower) Cross River, Benue-Congo)	Chapter 18
Otoro (Heiban, Kordofanian)	Chapter 19
Supyire (Senufic)	Chapter 20
Yoruba (Volta-Niger)	Chapter 21
Zande (Ubangi, Gur-Adamawa-Ubangi))	Chapter 22

0.3 The structure of the book and of the chapters

As mentioned above, there is little broad public knowledge of the details of Niger-Congo languages, either of the verb or any other component. Just because the phylum is so huge, scholars of Niger-Congo have tended to concentrate on one or two languages, one family, or one small part. They know their selected area well but usually have a restricted, passive knowledge of other families in the phylum. Large areas are under-described or not described at all, as can be seen by consulting the World Atlas of Language Features Online. And while descriptions and analyses of better know languages are easily available, those for lesser-known languages are not always easily accessed. So non-Africanists know even less than the specialists.

We attempt to address this problem by devoting one chapter to each of our twenty-one chosen languages. Our goal in each chapter is to present a basic description of facts and data while at the same time demonstrating our aims, methods, and assumptions. Our hope is that readers will be able to read each chapter as a self-standing description of its language. The source materials are quite diverse, being of different length and having different purposes and theoretical frameworks. To make the task of readers easier, we have taken two steps. One is to have a more or less standard format for each of the language chapters: section 1, 'General', presents general facts about the community speaking the language, and other bits of information

we deemed relevant or interesting. Unless otherwise stated our source for population size is Gordon 2005; section 2, 'Word Order' has a typological statement about basic word order, plus variants, if we were aware of them; section 3, 'Verb Structure', shows a verb template and inventory of grammatical morphemes for each position in the template; section 4, 'Tense and Aspect', or just 'Aspect', sets out the facts and our analysis; section 5, 'Other categories', deals with mood, focus, imperatives, and relativization, because we saw that these categories play an important role in the verb. For some languages, coverage of these categories is less than complete, because we could not find details in the source. Section 6 discusses briefly negation, a central verbal feature. Finally, section 7 offers data on auxiliary verbs, especially 'be' verbs, because they also play an important role in verbal categories and often illustrate grammaticalization in process. Not every chapter has all of these sections in this detail, but they all attempt to deal with as many of these features as the source allowed.

The second step is that we have taken some liberties with the content of sources. Most were written in the late twentieth century in different conceptual frameworks – generative, functional, "traditional", descriptive - and they use a range of terminology. We ourselves found it difficult in some cases to unravel it, and so we have presented everything in a more or less unified framework. This makes for easier reading, even though it might upset some of the source authors. We point out where we depart from our sources. We have been in contact with many authors to get their reaction to our procedure and approach. We had hoped to use a more or less unified terminology set out in one chapter but in the end this proved beyond us. While all three authors shared many theoretical assumptions, we disagreed strongly on others, and we could not arrive at a unified approach nor a unified terminology. The chapters by Hewson use one set of terminology (explained in the material after Chapter 22), those by Rose use another set (explained in her Brief Note on Terminology), and those by Nurse use a third set (more or less explained in Chapter 1).

Each chapter concludes with a Bibliography of works consulted for that chapter. Generally used abbreviations appear in the List of Abbreviations. However, because sources for the various languages often use terms and/or abbreviations specific to their language, in such cases we have indicated this at the start of individual chapters.

0.4 Morphosyntactic background

We assume that many readers will not be familiar with the typology of Niger-Congo languages so this section includes an outline of word order, verb structures, and serial verbs. These issues are mentioned in each of the individual chapters.

0.4.1 Word order in Niger-Congo

Verb categories in Niger-Congo are linked to verb morphology, which in turn is linked to the order of sentence constituents, so we start by sketching word order and verb structure. The sketch combines typology and diachrony.

Greenberg's classic word order typology (1963) distinguished three main orders for sentence constituents, referred to as SVO, SOV, and VSO. Heine (1976) added a fourth, giving the current standard framework for talking about constituent word order in African languages. The three types shared by Greenberg and Heine are labeled A (SVO), C (VSO), and D (SOV) in what follows, and Heine's additional type is B.

In (1), the abbreviation AUX is a label for a slot, which might contain auxiliary verbs, particles, or adverbials. Since AUX is the label for all the material between S and V (or O), it refers to quite different material in synthetic versus analytic languages. "X" refers to 'other' constituents (see Abbreviations).

(1) **Word order in Niger-Congo**

Language (and family)	Basic word order	Other
Aghem (Grassfields Bantu)	S (AUX)V O X	Some Grassfields have SAUXOV (L. Hyman p.c.)
Bambara (Mande)	S (AUX) O V X	
(Narrow) Bantu	S (AUX) V O X	One language has OV (Mous 2005). Many have SOV if O is a pronoun
Bijago (Atlantic)	S (AUX) V O X	SOV when O is 1 or 2 pl pronoun
Degema (Volta-Niger)	S (AUX) V O X	
Dogon	S X O V (AUX)	SVO in e.g. relative clauses
Doyayo (Adamawa)	S (AUX) V O X	SOV with cognate objects. Some Adamawa lgs have basic SOV
Ejagham (Ekoid)	S (AUX) V O X	
Ewe (Kwa)	S AUX V O X	SAUXOV-Nom in PRG/PRS
Fula (N. Atlantic)	S (AUX) V O X	VSO in some subordinate clauses (e.g. relatives)
Godie (Kru)	S V O X	SAUXOV-Nom
Ijo (Ijoid)	S X O V (AUX)	
Jukun (Benue-Congo)	S V O X	SAUXOV-Nom
Kabiye Gur)	S V O X	SAUXV-Nom O
Kisi (S. Atlantic)	S V O X	SAUXOV (does V = V-Nom?)
Makaa (Narrow Bantu)	S V O X	
Obolo (Cross River)	S V O X	
Otoro	S V O X	Other Kordofanian lgs have other orders (SOVAUX, even VSO)
Supyire (Senufic)	S AUX O V X	
Yoruba (Volta-Niger)	S AUX V O X	Also SAUXXVO ?
Zande (Ubangi)	S V O X	VSO in some subordinate clauses

Type A languages have a basic S V O X order, and prepositions[2]. Within the noun phrase there is some variation, most languages having all nominal modifiers after the head noun, but a minority having the adjectival modifier before, with all other nominal modifiers after, the head noun[3]. This is the commonest order worldwide, in Africa (Heine 1976, "71%"), and in our Niger-Congo sample (13 of 21).

[2]"Basic" refers to word order where S and O are nominals (in some languages they may also be pronominals). Type A languages may or may not have an AUX before V, thus S AUX VO X.

[3]As Creissels (2000:253) points out, word order in the noun phrase tends to harmonize with clause constituent order: in VO languages, as the verb comes before its complements, so the noun precedes its modifiers, while in OV languages, verb follows complement and noun follows its modifiers.

As Type A, Type C languages have prepositions and nominal modifiers following the noun but differ from A in having the verb before the S, thus V S O X. Type C is uncommon worldwide and in Africa (Heine 1976, 5%), and largely irrelevant to Niger-Congo, reported as occurring as the basic constituent order only in a few Kordofanian languages not part of our sample, and occurring in a few languages in our sample only as the word order in some subordinate clauses (e.g., Fula, Zande). Its presence in some Kordofanian languages likely results from contact[4].

All Type D languages have SOV and postpositions but there are two subtypes: in one type all nominal modifiers precede the head noun, whereas the other has nominal modifiers after the head noun. For Creissels (2000:252) "true" SOV languages have the AUX following the V. So in our sample only Ijo and Dogon would be "true' SOV languages, while Bambara and Supyire would not be, because they have AUX before the V[5]. Type D languages are not widespread in Africa and where they do occur they usually alternate with one of the other orders, most often SVO, in some functions.

Type B languages may have SVOX or SOVX or both. What they have in common is postpositions, and genitival modifiers before the head noun but all other nominal modifiers following the head noun. Since Type B languages are rare outside Africa but fairly widespread in Africa, especially West and Central Africa (see Güldemann 2008:159-63[6]), this requires an explanation. The current most widely accepted solution is that of Heine & Claudi (2001:43), who proposed that the co-occurrence of the SVO order with GEN-N order in the noun phrase is responsible for the emergence of the S AUX O V order via a nominalization process, described thus by Creissels (2000:241):

> ..in such languages (having SVO, and genitive modifiers before the head noun), noun phrases corresponding to the object of finite verb forms precede nominalised verb forms, since they are treated as their genitival modifier; consequently, with complex verb forms consisting of an auxiliary verb and a nominalised form of the main verb, the noun phrase corresponding to the object of a finite verb form precedes the nominalised form of the main verb. Subsequently, the decategorialisation of the auxiliary verb leads to the reanalysis of such constructions as involving a finite verb form preceded by a noun phrase in object function.

So in this scenario, SVO with GEN-N leads to S AUX OV-Nom and ultimately to SOV when the auxiliary is decategorised. In our sample, beside a majority (13) with SVO, a minority (4) has both SVO and S AUX OV[7], and as we have just seen, another minority (4) has only SOV today. Creissels (2000:241) also says: "The presence of TAM markers between S and O in clauses with the constituent order SOVX can be viewed as a strong hint that such a process of reanalysis took place in the history of the language in question". Today some Grassfields languages (Bambara (Mande), Ewe (Kwa), Godie (Kru), Jukun (Benue-Congo), Kisi (S.

[4]There is more typological variation within Kordofanian than any other NC family.

[5]Watters (2000:199) says S AUX OV is "more common" in Africa than S O V AUX.

[6]The co-existence of S V O and S O V cuts across genetic boundaries in Africa, occurring not just in Niger-Congo but also in members of the three other phyla, Afro-Asiatic, Nilo-Saharan, and Khoisan.

[7]This may be an understatement because not all our groups behave unanimously. Thus Aghem, our Grassfields Bantu representative, has S V O, but other Grassfields languages have S V O and S AUX OV (Larry Hyman, p.c.).

Atlantic), Supyire (Senufo), and, according to Güldemann (2008:261) some Adamawa-Ubangi languages) have OV with auxiliaries and/or TAM markers before the O.

If this scenario is correct, when did SOV emerge from SVO? The weight of recent publications[8] is on Proto-Niger-Congo (and Proto-Bantu) having S AUX O V beside basic SVO. That Proto-Niger-Congo had an analytic structure in which a string of discrete items, including auxiliaries (AUX), preceded the V is not in dispute, so S AUX V O. While it is possible that S AUX O V might have developed out of this already at the proto-stage, it seems likely that the conditions that gave rise to the nominalization continued to exist long after the proto-language. If they continued to exist, then later grammaticalizations of the same kind could have continued to arise during the later development of Niger-Congo. Indeed several of the languages examined still have a nominalised V, which suggests development from a more recent grammaticalization, rather than one that characterized the proto-languages ten millennia ago. The details need further analysis.

Basic word order can be modified by any of several factors: (i) One is topicalization or focalization, which moves the highlighted constituent out of its regular position; (ii) A word order other than the basic one often occurs in subordinate clauses, especially relative clauses: or where the object is a pronoun, not a noun; (iii) In some languages, one order is associated with certain aspects/tenses, while another co-occurs with other aspects/tenses, which might be linked to the chronology of the grammaticalisation involved; (iv) Finally, while word order is clearly linked to genetic inheritance and to these internal factors, it can also be modified by areal diffusion. The greatest variation is in Kordofanian, which has stood at the crossroads between representatives of the different phyla for millennia, while in West and Central Africa, the S AUX O V order of Type B occurs in Niger-Congo, Nilo-Saharan, and Afro-Asiatic languages.

0.4.2 The structure(s) of verbs in Niger-Congo

As can be seen in (2), many Niger-Congo families (11 of 21) have a nuclear stem, consisting of Root – Extension (EXT) - Final Vowel (FV), in which EXT and FV are affixes. The stem or what remains of it is bracketed in (2). EXT includes morphemes for valency-changing categories such as causative, applicative, neuter, reciprocal, reversive, and passive[9]. FV was originally used for a binary aspect contrast between perfective/factative and imperfective, both indicated by a single vowel. Although various groups now use this position to indicate other categories, such as subjunctive, it is not clear from the NC evidence that it was originally so used. We can assume this affixal stem structure goes back to Proto-Niger-Congo (Nurse 2007), because it is very unlikely that so many geographically distant families would have innovated it independently.

Eleven of twenty one families have an EXT, though the size of the set at EXT varies from language to language. While most members of most sets are shared, suggesting they are old and inherited, some are apparently innovations. Six (Bambara, Ewe, Ijo, Jukun, Obolo, Yoruba) have no trace of an EXT, although Williamson & Blench (2000:30,33) say of Ewe and Obolo that other languages in their families do have EXTs. Three families (Grassfields, Ekoid, Senufic) have rolled traces of the EXT into a single post-radical portmanteau position, labeled Suffix in (2).

[8]Gensler (1994, 1997), Gensler & Güldemann (2003), Güldemann (2008:159-62), Creissels et al (2008:127).

[9]Although Welmers (1973) says only Bantu shows a clear passive within Niger-Congo our survey showed other groups with a passive, though not of the Bantu shape.

Eleven of twenty one have the discrete FV slot. Four (Aghem, Kisi, Obolo, Yoruba) have no trace of the FV while six (Bambara, Ejagham, Ewe, Ijo, Jukun, Supyire) have a Suffix, which rolls together some former EXT morphemes and some from FV[10].

(2) **Verb structures in Niger-Congo** (Nurse 2007)[11]

Language	Verb structure
Aghem	HYP SP CFL NEG T-FOC SBJ [root - suffix] HAB NEG FOC
Bambara	SP AUX (includes A, NEG, T) CAUS – [root - suffix]
Bantu	PreSM – SM – NEG – TAM – OM – [root - EXT - FV] - IMP Pl
Bijago	NEG – FOC – SM – A –"AM" – OM – [root - EXT - FV] - REL
Degema	(1) SM – [root - EXT = FV] – IMP NEG (2) SM – [root = A] SM - root
Dogon	(1) [root – EXT – FV] – NEG – TAM – SM (2) [root – TAM] [AUX – NEG – TAM – SM]
Doyayo	SP AUX [root - EXT - FV]
Ejagham	SM – NEG/A – REP – [root - suffix]
Ewe	SP = NEG = M_1 = A DIR_1 M_2 M_3 DIR_2 "be' [root – suffix] = OP NEG
Fula	NEG SP - AUX - [root – EXT – FV] – NEG – no – IOM – DOM – LOC/INST
Godie	SP – a [root – EXT – OM – FV]
Ijo	SP OP – [root – EXT – $suffix_1$ – $suffix_2$] AUX
Jukun	SM – NEG – AM – [root – suffix]
Kabiye	SM – NEG – A – [root – EXT – FV] – OP
Kisi	SP [root – EXT]
Makaa	SP T H_1 P1 NEG=CM HAB PRG ADV AUX OM – [root – EXT – FV – H_2]
Obolo	M – SM – NEG – AM – [root]
Otoro	na – HAB – SM – TA – [root – EXT – FV] – OP – IMP
Supyire	SP NEG AUX OP IP – FP – [root – suffix]
Yoruba	SP AUX [root]
Zande	SM (or SP?) – TAM – [root – EXT – FV] – OP – NEG

The original nuclear structure Root – EXT - Final Vowel has been lost in languages in a large area of West Africa: northwestern Bantu, Grassfields Bantu, many Bantoid languages, and languages farther west, such as Yoruba. Eight families have only a Suffix, which incorporates remnants of EXT and FV, besides some new material, while two (Yoruba, Obolo) have nothing after the root.

Several processes seem to be involved in this loss. First, the derivational extensions at EXT could no longer be expressed because the prosodic stem became limited to four, three, and then two syllables (Hyman 2004), affecting material at the right of the stem. Then, although segmental material was deleted, associated tones were not, leading to the appearance of floating tones, often associated with the expression of aspectual categories, and to an expanded role for

[10]FV and Suffix differ in this text: FV follows EXT, but Suffix is EXT and FV rolled together.

[11]Items separated by a space are discrete: the equals sign represents cliticisation, where known, and the hyphen indicates affixation. Abbreviations are listed in the Abbreviation and Conventions section.

tones in general. If a language loses the segmental expression of derivational and aspectual categories, in principle it faces a choice: it can also lose the categories, or it can express them some other way. In practice, that seems to be hardly a choice as all the languages looked at that have lost extensions and/or final vowels, keep the categories and express them some other way. Typically, for example, the functions of the applicative extension ("to, from, for") are replaced by the use of word order, or prepositions, while the causative is expressed by use of some auxiliary, "to make/cause to verb", as can be seen illustrated in the first Ejagham example in (3a). The expanded role of tones can be seen examples (3b-e), which differ only tonally. Care is necessary here, as it is not clear whether the surface tones in these examples reflect floating tones, or tonal patterns associated with individual aspects and moods (and elsewhere, tenses), or come from prefixes or the stem itself.

(3) **Examples of replacement of EXTs and FVs, in Ejagham** (Watters 1981)

a CAU **a-Ø-yím** **ɛtá** **a-Ø-gbɔ̂** (verb **-yim** 'make, do')
3s-Ø-make Eta 3s-Ø-fall
'He made Eta fall.'

b PFT **a-Ø-fag** 'They have swept' (verb **fag** 'sweep')

c FAC **a-Ø-fâg** 'They swept'

d CON **á$^{!}$-Ø-fág** 'If they sweep'

e SBJ **á-Ø-fág** 'They should sweep'

While most Niger-Congo languages share the verb nucleus root - extension - final vowel, just outlined, they vary considerably in what precedes it. Current Niger-Congo languages range along a continuum between "analytic" and "synthetic" languages. In fully analytic languages what precedes the nucleus is a string of discrete pre-stem material (particles, auxiliaries, adverbials), indicating categories such as subject agreement, tense, aspect, mood, negation, focus, relativization, and often other categories such as conditional and object marking. Sometimes this pre-stem material is a long string, as implied in Ewe and Makaa templates in (2): sometimes the string contains serial verbs, often bare stems where most categories are only marked on one verb in the string, usually the first: sometimes an auxiliary (AUX in (2)) is followed by an infinitival main verb: sometimes the AUX is a form of 'be, or 'locative be', and the main verb is nominalized or locativized. Fully synthetic languages have fused all the pre-stem material, and so the nucleus is preceded by a set of inflectional prefixes. Between the ends of the continuum are many languages which have fused some morphemes while keeping others discrete.

Judgements about which families have a synthetic, and which an analytic structure are obscured in some cases by differences in how to interpret the data morphologically, but it appears that eight languages (Aghem, Bambara, Doyayo, Ewe, Kisi, Makaa, Supyire, Yoruba) are fully analytic, nine are synthetic (Narrow Bantu excluding the northwest, Bijago, Ejagham, Fula, Jukun, Kabiye, Obolo, Otoro, Zande), and three show some synthesis (Degema, Godie, Ijo). Dogon is excluded because all its inflection is post-stem.

There is a clear assumption that the analytic structure was original and that the grammaticalized structures developed from the analytic by cliticization and fusion. Of those Niger-Congo languages/families that have evolved a fully synthetic verb structure, some are adjacent or near Bantu: Ejagham (Ekoid, SE Nigeria/SW Cameroon), Jukun (Jukunoid, SE Nigeria), Obolo (Cross River, SE Nigeria), Zande (Ubangi, northern Democratic Republic of the Congo). Others are distant from Bantu: Bijago (Guinea Bissau), Fula (Senegal to Sudan), Kabiye (Togo), Otoro (southern Sudan). In the latter languages/families, synthetic structures resulting from grammaticalization are likely to have developed independently, whereas the emergence of synthetic structures in the families adjacent to Bantu suggests that they may be related to what happened in Bantu.

0.4.3 Serial verbs

While 'serial verbs' do occur outside Africa (Heine & Leyew 2008:22), they are particularly common in Africa, and especially in Niger-Congo (Heine & Leyew (ibid), Dimmendaal 2008:298)[12]. We have noted them for only eight languages but are aware that if we had searched more diligently, the number would have been larger as they are particularly common in parts of West Africa. Serial verbs can be defined loosely ("constructions in which two or more verbal lexemes combine without any overt indication of a dependency between them: none of the verbs is morphologically marked as dependent…no conjunction between them") or more narrowly ("constructions that involve two or more verbs but that, taken as a whole, have the behaviour of a single predicate, and not that of a construction involving distinct predicates in some dependency relation") (both from Creissels et al 2008:112).

(4) Yoruba[13] **ojó ra iwé fún iyá**
Ojo buy book give mother
'Ojo bought a book for mother.'

References

Creissels, Denis. 2000. Typology. In Heine & Nurse (eds.), 231-258.

Creissels, D, Gerrit J. Dimmendaal, Zygmunt Frajzyngier, C. König. 2008. Africa as a morphosyntactic area. In Heine & Nurse (eds.), 86-155.

Dimmendaal, Gerrit. 2001. Language shift and morphological convergence in the Nilotic area. In Nurse (ed.), 83-124.

---- 2008. Africa's verb-final languages. In Heine & Nurse (eds.), 272-308.

Gensler, Orin D. 1994. On reconstructing the syntagm S-Aux-O-V to Proto-Niger-Congo. Berkeley Linguistic Society (special session) 20:1-20.

---- 1997. Grammaticalization, typology, and Niger-Congo word order: progress on a still

[12] They occur outside Niger-Congo, for instance in Khoisan (Güldemann & Vossen (2000:109)).
[13] Example from Creissels et al (2008:113).

unresolved problem. JALL 18.1:57-93.

---- & T. Güldemann. 2003. S-Aux-O-V-other: typological and areal perspective. Presented at 4th WOCAL workshop on Distributed Predicative syntax (S P O V X). Rutgers, 6/2003.

Gordon, Raymond. G. 2005. *Ethnologue*. 15th edition. Dallas: SIL International.

Greenberg, Joseph.1983. Some areal characteristics of African languages. In *Current Approaches to African Linguistics 1*, ed. Ivan Dihoff, 3-21. The Netherlands: Dordrecht and Cinnaminson, N.J.: Foris Publications.

Güldemann. Tom. 2008. The Macro-Sudan Belt: towards identifying a linguistic area in northern sub-Saharan Africa. In Heine & Nurse (eds.), 151-85.

---- & Rainer Vossen. 2000. Khoisan. In Heine & Nurse (eds.), 99-122.

Heine, Bernd. 1976. *A typology of African languages based on the order of meaningful elements*. Berlin: Dietrich Reimer.

---- & Ulrike Claudi. 2001. On split word order: Explaining syntactic variation. *General Linguistics 38*, 1:41-74.

---- & Z. Leyew. 2008. Is Africa a linguistic area? In Heine & Nurse (eds.), 15-35.

---- & D. Nurse (eds.). 2000. *African languages*. Cambridge: Cambridge University Press.

---- & D. Nurse (eds.). 2008. *The linguistic geography of Africa*. Cambridge: Cambridge University Press.

Hyman, Larry. M. 2004. How to become a Kwa verb. *Journal of West African Languages* 30.69-88.

Mous, Maarten. 2003. Nen (A44). In Nurse & Philippson (eds.), 283-306.

----- 2005. The innovative character of Object - Verb word order in Nen (Bantu A44, Cameroon). In Bostoen, K. & J. Maniacky (eds.). *Studies in African comparative Linguistics with special focus on Bantu and Mande; essays in honour of Y. Bastin and C. Grégoire* (pp. 411-424). Tervuren: Koninklijk Museum Midden Afrika.

Nurse, Derek. 2007. Did the Proto-Bantu verb have an analytic or a synthetic structure? In Kula, N.C. & L. Marten (eds.). *Bantu in Bloomsbury. Special Issue of the SOAS Working Papers in Linguistics*. Volume 15. London: Department of Linguistics SOAS, 239-256.

---- 2008. *Tense and Aspect in Bantu*. Oxford: Oxford University Press.

Watters, John. R. 1979. Focus in Aghem. In Hyman (ed.), 137-198.

---- 1981. *A Phonology and Morphology of Ejagham, with notes on dialect variation*. PhD thesis. University of California at Los Angeles.

---- 2000. Syntax. In Heine & Nurse (eds.), 194-230.

Welmers, William E. 1973. *African Language Structures*. Berkeley: University of California Press.