

West African languages

Linguistic theory
and communication



Edited by
Nina Pawlak
Izabela Will



West African languages

West African languages

**Linguistic theory
and communication**

**Edited by
Nina Pawlak
Izabela Will**



Reviewers

Jadwiga Linde-Usiekniewicz

Anne Storch

Commissioning Editor

Ewa Wyszynska

Proofreading

Bożena Lesiuk

Cover Design

Anna Gogolewska

Illustration on the Cover

siloto/shutterstock

Layout and Typesetting

Dariusz Górski

Published with financial support from the University of Warsaw

Published with financial support from the Faculty of Oriental Studies, University of Warsaw

Published with financial support from the University of Warsaw Foundation



© Copyright by Wydawnictwa Uniwersytetu Warszawskiego, Warszawa 2020

Nina Pawlak ORCID 0000-0001-6949-3921

Izabela Will ORCID 0000-0002-2797-1160

ISBN 978-83-235-4623-8 (print)

ISBN 978-83-235-4631-3 (pdf online)

ISBN 978-83-235-4639-9 (e-pub)

ISBN 978-83-235-4647-4 (mobi)

Wydawnictwa Uniwersytetu Warszawskiego

00-838 Warszawa, ul. Prosta 69

E-mail: wuw@uw.edu.pl

Online bookstore: www.wuw.pl

1st Edition, Warsaw 2020

Print and binding

POZKAL

■ Contents

General abbreviations and symbols	7
Standard glossing labels	7
Introduction <i>Nina Pawlak, Izabela Will.</i>	9
PART I. WEST AFRICA AS A LINGUISTIC AREA	
Measuring phonological complexity in West African languages <i>Gian Claudio Batic</i>	23
Innovative features of nouns and pronouns in Chadic languages of the Nigerian Gongola-Benue basin <i>Sergio Baldi, Rudolf Leger.</i>	46
Methodological and technical challenges of a corpus-based study of Nijja <i>Bernard Caron</i>	57
PART II. STRUCTURAL PROPERTIES OF LANGUAGES AND THEIR THEORETICAL FRAMES	
Clause chaining in Bambara <i>Valentin Vydrin</i>	79
Categorization of phasal polarity items in Bambara (Mande) <i>Klaudia Dombrovsky-Hahn</i>	99
Ma-causatives in Tomo Kan Dogon: between causatives and passives <i>Vadim Dyachkov</i>	127
Asymmetries in negation in the Atlantic languages: a preliminary comparison <i>Aurore Montébran</i>	140
Let's tidy up the grammar of Hausa and analyse <i>after/behind</i> in <i>after class/after he arrives/he's behind</i> etc. as prepositions <i>Philip J. Jaggard</i>	153
Enhancing grammar and valence resources for Akan and Ga <i>Dorothee Beermann, Lars Hellan</i>	166
PART III. SYNTACTIC STRUCTURES AND CONCEPTUALIZATION PATTERNS	
On the <i>shi-</i> evaluative prefix in Lamnso' <i>Lenzemo Constantine Yuka.</i>	189

Tiv compounds and compounding <i>Vanessa Chivir Adzer</i>	198
Nigerian Pidgin English phraseology in the context of areal influences <i>Olga Frąckiewicz</i>	213
Lexical iconicity in Adamorobe Sign Language (AdaSL) and Ghanaian Sign Language (GSL) <i>Mary Edward</i>	230

PART IV. SOCIAL AND COMMUNICATIVE ASPECTS OF LANGUAGES

Phraseological units involving body-part terms: a corpus based analysis of Hausa to English translation <i>Yakubu Magaji Azare</i>	255
Linguistic repertoires of refugees in Internally Displaced Peoples' (IDP) camps in North East Nigeria <i>Judith A. Mgbemena</i>	266
The function of nonverbal regulators in Hausa face-to-face interaction <i>Aisha Umar Adamu</i>	279
“Allah expressions” as a manifestation of common cultural area in West Africa <i>Nina Pawlak</i>	288

General abbreviations and symbols

C	Consonant	PP	Prepositional Phrase
H	High tone	SVC	Serial verb construction
L	Low tone	TAM	Tense-Aspect-Mood
M	Mid tone	WALS	<i>World Atlas of Language Structures</i>
NP	Noun Phrase	V	Vowel
POS	Part of Speech		

Standard glossing labels

1	first person	GEN	genitive
2	second person	ICP	intransitive copy pronoun
3	third person	IDEO	ideophone
4	impersonal pronoun	IO	indirect object
ACC	accusative	IMP	imperative
ADJ	Adjective	IPFV	imperfective
ADV	adverb(ial)	LOC	locative
AGR	agreement	M	masculine
AUX	auxiliary	NEG	negation, negative
COMP	complementizer	OBJ	object
COMPL	completive	PART	particle
CONT	continuous	PFV	perfective
COP	copula	PL	plural
DAT	dative	PLUR	pluractional (verbal plural)
DEF	definite	POSS	possessive
DEM	demonstrative	PRS	present
DET	determiner	Q	question
F	feminine	REL	relative
FOC	focus	SBJV	subjunctive
FUT	future	SG	singular

Nina Pawlak

ORCID 0000-0001-6949-3921

University of Warsaw

Izabela Will

ORCID 0000-0002-2797-1160

University of Warsaw

■ Introduction

In the more than 150-year-old tradition of studying African languages, West Africa has become an area of many research developments related to individual languages or language groups and families, which, nevertheless, represent particular research circles rather than a single branch of African linguistics.¹ Various scholars and institutions contributed to the knowledge of the structural peculiarities of languages with different affiliations to language families (Afroasiatic, Nilosaharan and Niger-Congo) that in West Africa are used within relatively small distances. Research on areal phenomena in African linguistic studies has begun rather late (Aikhenvald & Dixon 2001, Heine & Kuteva 2001, Dimmendahl 2008, 2011), just like studies on “contact zones and linguistic convergence areas in West Africa” (Cyffer & Ziegelmeyer, ed. 2009; Zima 2009, among others) which remain a subject of constant discussion. The issue of contact zones includes a wider reflection on the interference of genetic and areal factors as well as on linguistic tools and methodological background of comparative works, which comprise a variety of structural patterns and different attitudes to their interpretation. The descriptive works are often combined with sociolinguistic and communicative studies on languages, which set many new research directions and give them a more unified perspective.

The present volume joins two traditions, i.e. research conducted in separate linguistic circles with challenges of looking at West African languages from a broader areal perspective. It contributes to the studies on structural properties of the selected languages and corresponding methodological approaches to their analysis, but also provides some new research findings related to social and communicative aspects of the languages in West African context. Such an attitude results from interaction and exchange of ideas among scholars of African languages experienced in Chadic (or Afroasiatic) linguistics, Mande studies, Atlantic studies, and Akan studies or various working groups, dedicated to specific areas of language research.

The idea of the volume emanates from presentations and discussions conducted during the series of conferences known as Symposium on West African Languages, initiated in Naples in 2014 and continued in Vienna in 2016. Most topics present in the volume were discussed during the 3rd edition of the conference held at the University of Warsaw in 2018 (SyWAL2018).

The authors represent main centers of African studies in which West African languages are subject of research. These include universities and research institutions from Europe

¹ On comparative and descriptive African linguistics see Part II in Wolff. 2019: 137-416

(Paris, Naples, London, Brighton, Bayreuth, Frankfurt am Main, Moscow, Trondheim, Warsaw), and Nigeria (Kano, Benin, Ibadan, Makurdi, Wukari). The cooperation on institutional ground promotes exchange of views between the previously established research circles and evokes the discussion on new ideas that arise from the current state of research on West African languages.

The contributors discuss properties of West African languages having a long descriptive tradition (Hausa) and those that are subject of intensive studies in recent decades (Bambara), as well as features of less studied or endangered languages (Tiv, Lamnso', Ga, Tomo Kan). The comparative analyses are either genetic and limited to language groups (Bole-Tangale, Atlantic languages) or have an areal perspective to show their wider West African dimension. Linguistic processes are discussed along with issues concerning communication and sociolinguistics, making the relation between language and society clear and documenting the association of distinctive features of West African languages with cultural values.

1. West Africa as an area of linguistic diversity and unification processes

West Africa as a sub-region of the African continent is defined mostly on geographic and political criteria which exclude Northern Africa and the Maghreb covering the Sub-Saharan countries from Senegal to Nigeria. As a region of linguistic studies, West Africa adheres to these limits, though genetic relationship and historical contacts between languages make these conventional boundaries vague in a number of respects. The region is characterized by linguistic diversity which determines the prominence of research oriented at multilingualism and language contact. The works conducted so far has focused on identifying convergence zones rather than providing the proof of the linguistic coherence in the entire region. The term convergence zone refers to a region where many linguistic features are shared across the language boundaries. The two largest units, i.e. Macro-Sudan Belt extending from Senegal to Ethiopia (Güldemann 2008) and Wider Lake Chad Region overlap to some extent, especially in Nigeria where genetically distinct and structurally different languages meet (Ziegelmeyer 2015; Wolff & Löhr 2005; Zima 2009; Cyffer & Ziegelmeyer 2009).

West Africa is characterized by extensive societal multilingualism (Lüpke & Chambers 2010). Along with indigenous languages superimposed foreign languages such as French, English or Portuguese are used. The region has always been an area where languages brought by scholars, traders or travelers were in constant confrontation with those used locally. Muslim teachers and traders moving along West Africa brought Arabic to this region. The emergence of political centers such as Ghana in 12th century, Mali in 14th century, Songhay in 15th century or the Sokoto Caliphate in 19th century strengthened the position of Arabic as the language of courts, written correspondence, religious and legal teaching. Arabic also became an important contact language among educated and influential people living in towns and it had an impact on the major languages of the Sahel and northern savannah. The historical empires and city-states also promoted languages spoken by the ruling class such as Hausa, Fulfulde or Mande languages, pushing many other local languages aside. Due to globalization, urbanization and economic development, the number of languages

spoken not only in the entire macro-region, but also in particular towns or districts is decreasing. However, there are still places where the extensive multilingualism is the daily bread. The diversity is particularly prominent in the Nigerian Middle Belt, where the minority languages function in a wide range of circumstances. It is also present in Casamance, southern region of Senegal squished between Gambia and Guinea-Bissau, where more than ten languages are spoken. Several lingua francas of the region are: Mandinka, Peul, Joola Fogany, French, and Portuguese Creole. Other languages such as Balanta, Manjaku, and smaller Joola languages serve as lingua francas on a smaller scale. In addition, minority languages such as Bainounk cluster or Bayot are used. The national language of Senegal – Wolof, which has been spreading rapidly since the country’s independence in 1960, also made its presence in the region (Cobbinah 2010: 177).

West Africa is also perceived as an area of linguistic change mediated by social factors (Dimmendaal 2008: 855) which affect languages in their sociolinguistic dimension. Looking from such perspective, multilingualism is perceived as a threat to existence of minority languages (Blench & McGill 2012) which become endangered. However, the multilingual situation is also an opportunity to spread the linguistic patterns characteristic of the dominant languages. Hausa remains an important source of grammatical structures and linguistic patterns shared by many languages of the region. The status of Hausa in West Africa and a long tradition of research on its structural features results in the availability of data for the extensive study of the grammatical system and study of the language situation in West Africa through the lens of Hausa language data. In more recent times, it is Pidgin English which gains popularity not only in Nigeria, but also in other parts of West Africa. Studies of its structural peculiarities provide insight into the most common patterns and ways of conceptualization characteristic of the West African region (Frackiewicz 2019). They clearly demonstrate the ongoing process of unification manifested by the use of the same structural patterns and structural calques in many languages. Such is the instance of complementizer in a number of West African languages, which has the form based on the verb ‘say’. It is worth noting that in Nigerian Pidgin English it is also *say* not *that* as in English, the lexifier language.

2. The structure of the volume

The volume is divided into four parts devoted to particular aspects relevant to general linguistic studies and studies on West African languages. Part I: **West Africa as a linguistic area** presents the results of comparative studies on languages, which refer to the prominent achievements of African linguistics in phonology and diachronic studies. It also includes a corpus-based study providing methods for linguistic analysis of unstable and changing structures. In the first article, “Measuring phonological complexity in West African languages” Gian Claudio Batic summarizes the areal aspects of phonological complexity and provides evidence that West Africa can be considered a coherent phonological area. Therefore, the idea which has been already postulated in the literature and identified on the map of phonological areas in Africa (Clements and Rialland 2008) has gained the support. Another article by Sergio Baldi and Rudolf Leger “Innovative features of nouns and pronouns in Chadic languages of the Nigerian Gongola-Benue basin” refers to diachronic

investigations on Chadic. The analysis goes beyond the traditional approach and leads to identifying external influences which affect grammar. They are motivated by contact with neighbouring Adamawa and Jarawan Bantu languages.

In “Methodological and technical challenges of a corpus-based study of Naija” Bernard Caron refers to a large corpus created within a long-term project. The article provides comprehensive information on the development and functional status of Common Nigerian Pidgin (or Naija), which functions as *lingua franca* in in Nigeria. This corpus, the first for creole and pidgin studies, creates a basis for the development of language and standardization of the orthography, which requires special methodological choices in annotation to establish the link between the change in structure and the change in language use and function.

Part II: **Structural properties of languages and their theoretical frames** deals with the interpretation of structural patterns characteristic of West African languages. It includes the topics corresponding to the descriptive tradition of main West African language groups, namely Mande, Atlantic, Chadic, and Kwa languages. Valentin Vydrin in “Clause chaining in Bambara” presents a clause linking type which is distinct from both coordination and subordination. Its nature is close to coordination syntactically, and resembles subordination morphologically. Providing a rich documentation of the construction labeled a “cosubordination” construction, the author contributes to a better understanding of mono- and pluriclausality in linguistic theory. Klaudia Dombrovsky-Hahn discusses “Categorization of phasal polarity items in Bambara (Mande)”. In Bambara, the markers of ALREADY, NOT YET, STILL and NO LONGER are classified as adverbs or particles. The analysis leads to identifying their lexical sources which are found in verbs and shows how their status has been changed in the process of grammaticalization. The data for the analysis enable presentation of some structural properties of Bambara, such as the feature of transitivity and intransitivity and a variety of exponents to express the meaning ‘to be’ with their positive and negative lexical counterparts. In the following article, “Ma-causatives in Tomo Kan Dogon: Between Causatives and Passives” Vadim Dyachkov discusses the causative/passive polysemy in Tomo Kan, a Dogon language (Niger-Congo) spoken in Mali and Burkina Faso. By using syntactic tests (modification by temporal adverbs, ellipsis, reflective binding), Dyachkov shows that distant causatives formed by adding the *má-* suffix are monoclausal structures where the causing and the caused sub-event cannot be modified independently. On the basis of structural properties of distant causatives (*má-* suffix and lack of subject marker), it is claimed that the emergence of passive meaning is associated with the loss of the clausal suffix (resulting in the absence of the overt subject) rather than in terms of reflexivization or pragmatic reinterpretation.

Structural properties of languages and their theoretical implications are discussed in the following three articles. In “Overview of asymmetries in negation in the Atlantic languages” Aurore Montébrun shows that negation is a complex process involving different grammatical categories, such as person and number marking, as well as the syntactic organisation of the sentence. There are more irregularities and asymmetries between affirmative and negative sentences than the presence or the absence of the negation marker. For example, completely different TAM markers or copulas are used in negative and affirmative clauses. The heterogeneity of the negative strategies and their interactions in Atlantic languages shed more light on internal clustering of the Atlantic group.

In his article “Let’s tidy up the grammar of Hausa and analyze ‘after/behind’ in ‘*after class/after he arrives/he’s behind*’ etc. as prepositions” Phillip Jaggard tries to re-examine the existing descriptions of Hausa prepositions from theoretical perspective not used before. He provides evidence for a unified analysis of prepositions in terms of heads of phrases and their complements. The idea has its roots in early work by Jespersen (1909-49), extended in Pullum & Huddleston (2002) with reference to English. Reinterpretation of the category of Hausa prepositions takes into account the morphological and distributional criteria. Though controversial in the context of Hausa descriptive tradition, it joins together all variants and alternative interpretations of the phrasal constituents which very often have the same lexical base.

Contributing to another field of West African studies, Dorothee Beermann and Lars Hellan in “Enhancing grammar and valence resources for Akan and Ga” present data management and analysis software for endangered languages. They demonstrate how digital resources and their linguistic specification for one language are helpful in the development of resources for another language. The article provides interpretation of the notions of grammatical relations and valence in closely related languages which, nevertheless, show resource asymmetry. The comparison of *ba* having the lexical meaning ‘come’ in Akan and Ga in terms of syntactic environment and situation type explains the development of contemporary syntactic systems.

Part III: **Syntactic structures and conceptualization patterns** covers a wide range of strategies in coding the meaning, which have their motivation in pragmatic associations and culturally motivated patterns of conceptualization, along with iconic strategies petrified in sign languages. Lenzemo Constantine Yuka in his article “On the *shi-* evaluative prefix in Lamnso” presents the pragmatic values of the grammatical prefix qualified as noun class marker in systemic descriptions. The claim is that the *shi-* prefix in Lamnso’ connotes more than just the diminutive aspects of the nominal under qualification and very often portrays the speaker’s attitudinal evaluation of the noun. Therefore it is rather employed as a strategy of persuading an interlocutor to look at the object from a particular perspective.

The next two articles in this section deal with multiword expressions as new means of coding the meaning. Usually seen in terms of petrified phrases, they are presented as specific innovations which function in languages in the process of rapid development and without a stabilized linguistic standard. Vanessa Chivir Adzer in “Tiv compounds and compounding” presents both regular and specific rules of word combinations. The resulting form, which serves to convey a new message, can be classified as a noun, although the components of this multiword expression are nouns, adjectives, and verbs. Olga Frąckiewicz discusses “Nigerian Pidgin English phraseology in the context of areal influences”. The article presents various types of multiword expressions which are based on English words, but in terms of structural features and conceptualization patterns represent collocations based on native West African languages. It is shown that the most productive model for compounding is verb + noun combination. One of the verbs productively used in compounding, is *chop* ‘to eat’, e.g. *chop life* ‘to enjoy’, *chop maut* ‘to kiss’, *chop burger* ‘put on weight’. Another source domain of idiomatic expressions in Nigerian Pidgin are body-part terms, which have grammatical functions and are used to code various semantic concepts.

Topics presented in the volume also concern sign languages used in West Africa. Mary Edwards in “Lexical iconicity in Adamorobe Sign Language (AdaSL) and Ghanaian Sign

Language (GSL)” discusses preference for specific iconic strategies to name handheld tools such as key, hammer, comb or pen among signers of two languages specified in the title as well as rural and urban gesturers (i.e. hearing people who perform gestures) from Ghana. She notices that signers and gesturers exhibit systematic preference for iconic representation of tools choosing an action-based sign depicting how the object is held (handling strategy) or depicting features of the object (instrument strategy). Both GSL and AdaSL signers prefer the instrument strategy, while gesturers show a greater preference for handling strategy, although rural gesturers demonstrate a high preference for instrument almost comparable to their preference for handling, which may be connected with the fact that they live among the community of AdaSL speakers. The research shows that AdaSL, a rural sign language with a long history has higher consistency and agreement in the use of the same strategy for individual items than GSL, an urban sign language that emerged in 1957. In a summary, the author makes an observation that the preference for instrument strategy in naming handheld tools seems to emerge quickly in new sign languages.

Part IV: **Social and communicative aspects of languages** focuses on extra-linguistic factors determining the use of languages. In West African context, the role of vehicular languages, such as Hausa, in multiethnic communication is crucial. They affect the lexicon and structure of minority languages, being a source of borrowings and transmitting language behavior. Hausa is not only the language of broader communication, but also a language of written literature, both native and translated. Translated works of European or American writers are a rich source for analyzing systemic devices, stylistic forms and cultural norms.

The section starts with the article “Phraseological units involving body part terms: A corpus based analysis of Hausa to English translation” in which Yakubu Magaji Azare presents results of the analysis based on parallel texts from two different languages – Hausa and English. The comparison of structures involving body-part terms in the two texts from the perspective of translation shows significant differences between the status of body-part terms in the two languages due to different conceptual patterns and grammaticalization paths.

Judith Mgbemena in “Linguistic Repertoires of Refugees in Internally Displaced Peoples’ (IDP) Camps in North East Nigeria” shows a linguistic situation in Nigeria from the perspective of endangered languages. In reports and publications referring to the impact of terrorism and violent conflict on national development, the language component is usually ignored. In a similar vein, humanitarian groups helping the people in need focus on social, economic, religious or medical aid, leaving the problem of language and communication aside. The paper focuses on language use spectrum in two Internally Displaced Persons’ camps situated in east-north Nigeria showing how forced mobility resulting from violent attacks affects linguistic ecology of Nigeria. It is shown that indigenous languages are rarely used at the camps and the IDP people are exposed to other languages like Hausa and English. It poses a threat to the sustenance of the displaced languages not only locally, but also globally, as many of these languages are endangered.

The forms of direct communication involving gestures are discussed in “The function of nonverbal regulators in Hausa face-to-face interaction” by Aisha Umar Adamu. She analyses the use of nonverbal signals called regulators in Hausa social interactions taking place in schools and at wedding ceremonies. The research shows that various gestures and body postures act as conversational regulators that help to keep the flow of the conversation. Regulators mark the beginning or the end of the conversations (e.g. turning the head aside)

or function as turn-taking signals (e.g. inverting the mouth in an upside down “u” shape, akimbo, beard stroking). Depending on the communication context, regulators can function as substitutes for verbal expressions or as an emphasis of a verbal statement.

The final article of the volume deals with religious phrases of Arabic origin which are commonly used in West African languages. Nina Pawlak in “Allah expressions as a manifestation of common cultural area in West Africa” presents a set of established phrases which have phonological variants in Hausa, Kanuri, Songhay, Wolof, Bambara and many other West African languages. Reported in the dictionary as Arabic loans (Baldi 2008) they are fully integrated into the languages. They are used not only in religious context but also in everyday communication and have many pragmatic functions. The process of their secularization is illustrated with the examples extracted from spoken and written Hausa.

3. Empirical approaches in the context of West African linguistics

Although the book is divided into four parts, the topics touched upon in various parts correspond to each other. Structural, conceptual and communicative aspects of languages discussed in parts II-IV often reflect phenomena common for the entire region. They refer to a well-documented fact that not only the lexicon, but also grammatical patterns like TAM systems (Wolff & Löhr 2005) and negation (Cyffr et al. 2009), can be borrowed. As far as structural patterns of negation in West Africa are concerned, they are often characterized by two factors: asymmetry and double marking. The asymmetric negation, i.e. the structure that exhibits more structural differences than the addition of a negative marker, is not only characteristic of Atlantic languages (as presented by Aurore Montebran in this volume). For example, Kanuri applies negative markers in the Completive and Potential TAMs, which is different from the Imperfect TAM. Negative Imperfect is derived from the construction where a negative predicator and an embedded affirmative clause were applied (e.g. the English sentence ‘I will not go to Kano’ was constructed as ‘there is not [that] I will go to Kano’; (Cyffer 2009: 1). As shown in *the Word Atlas of Language Structures*,² out of 22 West African languages that are represented in the data, 16 languages have both symmetric and asymmetric negation, 6 have only asymmetric negation while only in 2 of them negation is fully symmetric. Another feature characteristic of negation is double negation-marking – a syntactic structure that combines two different markers to express sentential negation. One of them is usually situated between the subject and the main verb, the other one is clause-final or sentence-final marker (Beyer 2009). Both features: double negation marking and asymmetry function in Lobi, another West African (Gur) language representing Niger-Congo family (Winkelmann & Mische 2009: 173). They confirm the areal representation of the specific patterns in non-related languages.

Another issue touched upon in papers dedicated to the description of particular languages, but having a more areal dimension is syntactic categorization. Classifying a given lexeme

² A large database of structural (phonological, grammatical, lexical) properties of languages (Dryer & Haspelmath 2013).

as a verb, adjective or preposition is often the result of descriptive traditions of a given language rather than the result of the formal analysis. The categorization is problematic due to the fact that many lexical items are multifunctional and one word such as *kòrɔ* in Bambara can function as the verb ‘get or be old’, the predicate adjective or state verb ‘old’, the modifying adjective ‘old’ or the noun ‘elder sibling’, among others (Dombrowsky-Hahn in this volume). As noticed by Caron (this volume) in order to avoid the problem of distinguishing verbs from adjectives in Nigerian Pidgin some researchers (e.g. Mazzoli 2013) use the term “property items” while others (e.g. Faraclaus 1989) state that the category “adjective” does not apply to the language; one can talk about “stative verbs” only.

Descriptions of West African languages are characterized by incoherence in the use of linguistic terminology, i.e. labels peculiar to the language structures. The articles of the volume contribute to the recognition of structural properties of languages, that may be helpful in unifying the terminology and doing comparative works with the use of compatible terms. This is exemplified by pluractionals, the term used widely in the description of Chadic verbal forms (Newman 2012), but not so common in the presentation of a very similar feature in other languages. Also, the term serial verb constructions may be subject of different theoretical interpretations and functional applications. Beermann and Hellan (this volume) notice that Ga (Kwa, Niger-Congo) makes use of multiverb expressions that subsume the so called Extended Verb Complexes, i.e. combinations of a main verb and verbs with a predominantly functional nature, like inducing deixis or path aspect. Such functional verbs precede the main verb (that is why they are often referred to as preverbs), but they can also occur post-verbally. In multiverb expressions the verb phrases may also play a role of adverbials. It shows that in areal attitude, there is a variety of methodological applications and the common basis of the interpretation is subject of further investigation.

The effect of language contact for grammatical structure finds its clear manifestation in the process of grammaticalization. Although cross-linguistically the whole African continent may be seen as one grammaticalization area (Heine & Kuteva 2015), West Africa is the region in which a set of the same conceptual schemas is used in grammatical functions, as for example, logophoric marking (Güldemann 2008) or exceed comparatives (Ziegelmeyer 2015). As shown by the authors of this volume, Philip Jaggard and Yakubu Azare, the semantic domain that functions as a rich source of grammatical markers are body part-terms. The grammaticalization of body-part terms results in the high frequency of their occurrence in texts and raises the question whether the given lexeme is still a noun or yet a pronoun or an adverb. The contributors to this volume opt for a formal solution of the problem and present various arguments towards such a claim. However, on a morphological level there is no overtly marked difference between the preposition and the noun. For example, in Hausa a single lexical form: *bayan* functions either as a noun ‘the back of (sth.)’ or a pronoun ‘at the back’.

The evidence of areal factors caused by multilingualism is most expressive in works on conceptualization patterns. The contributions go along with the statement that speakers using two languages continuously tend to minimize their cognitive and linguistic efforts to translate between these two languages. As a result, the speakers reorganize the range of meaning of words in each language so that they become closer, and reshape the way of structuring the events in the sentences (Ross 1997: 241). Multilingualism affects not only the lexical and structural part of the language. It also causes the emergence of common

conceptualization patterns. One of such patterns is defined as “African community model”, i.e. community model based on kinship (Polzenhagen 2007), characterized as the extension of the family concept to society in general (Polzenhagen 2007: 119). The semantic and conceptual extensions of kinship terms are based on metonymies, linking the kinship with the community, e.g. COMMUNITY FOR KINSHIP or KINSHIP FOR COMMUNITY (Hollington 2015: 170). For example, the term *baba*, which in Hausa and Yoruba means ‘father’, is used in Nigeria to refer to the elders in the community, to the leaders of various associations or even to the president of the country – *baba Buhari*. Another conceptualization pattern commonly used in West Africa is based on the extensive use of body-part terms. Not only are they used to refer to parts of every-day objects, but also to describe more abstract concepts like emotions, feelings or the process of thinking. Body-part terms are substantial part of idiomatic expressions or compounds in various languages spoken in West Africa. Conceptualization patterns find their exemplification in Hausa, e.g. *farin ciki* ‘happiness’ (lit. white stomach), *bakin ciki* ‘sadness’ (lit. black stomach), Pidgin English, e.g. *strong head* ‘stubborn, persistent’, *bodi du mi* ‘premonition’ (lit. body do me), *big ai* ‘ambitious, greedy’ (lit. big eye) and Tiv *shimányián* ‘anxiety (lit. red heart)’; (see the papers in this volume by Azare, Frąckiewicz and Adzer respectively)’. The contributors of the volume also confirm the feature of “semantic calquing” (Ziegelmeier 2015) which involves the verbs denoting “eating” and “drinking” in phrasal expressions (Frąckiewicz, this volume, see also Jaggar & Buba 2009).

The presentations of the volume are highly innovative when they discuss social and communicative aspects of languages, an emerging area of research in West African context. The issue of language domination in multilingual environment is presented for the Middle Belt region in Nigeria, which is characterized by the largest number of nation’s ethnic subsystems and language diversity and has a dominant local *lingua franca* – Hausa. It is shown (Mgbemena, this volume) how larger languages provide economic opportunities that minority languages do not offer. Speakers of major languages have an access to literature or text books, they can listen to radio or TV programs transmitted not only by local but also international broadcasting stations. The speakers of a larger language are privileged even in refugee camps, where they can get easier access to humanitarian and medical aid.

Some aspects of the studies on West African languages have a wider cross-linguistic dimension. Several conceptualization patterns and linguistic structures from West Africa are found in Caribbean languages. It is not possible to understand certain structures of Jamaican English such as serial verb constructions without looking at West African languages (Hollington 2015). The similarity is also evident in conceptualization patterns. As noticed by Ameka “[t]he connection between red eyes and envy and related emotions is not only an Ewe phenomenon but seems to be an areal feature in Ghana and West Africa (Ameka 2002: 29)”. The underlying conceptualization has been calqued into Ghanaian English, but also transferred to Caribbean languages. In Jamaican English the abstract concept of ‘envy’/‘jealousy’ is linked to ‘red eyes’ as in the expression: *di man dus tuu red yai* ‘the man is too envious’ (lit. the man has too red eyes; Hollington 2015:118). Due to mobility factor, the pidgin variety spoken in West Africa is not only spread locally but also internationally. It forged its path to Europe via broadcasting stations that have started offering news in Pidgin, recognizing its significance or to East Africa via Nigerian and Ghanaian films that are getting popularity in Kenya or Tanzania (Krings 2010).

The monograph presents salient features of West African languages and contributes to their interpretation in terms of typological distinctiveness. It is shown that the development of digital resources for West African studies and methods of their analysis allows comparison of similar structures in different languages and better understanding of their structural peculiarities. The discussion on particular features usually contains extended information about structural properties of the language, especially when they are less known in linguistic literature, and includes a note about its sociolinguistic status. The presentation vary in topics and methodological approaches, but the presented data provide evidence for identifying West Africa as an area of linguistic convergence and common cultural background.

References

- Aikhenvald, Alexandra Y. & Robert M. W. Dixon (eds.). 2001. *Areal diffusion and genetic inheritance: problems in comparative linguistics*. Oxford: Oxford University Press.
- Ameka, Felix K. 2002. "Cultural scripting of body parts for emotions. On 'jealousy' and related emotions in Ewe". In: Nick Enfield & Anna Wierzbicka (eds.). *The Body in Description of Emotion. Special Issue Pragmatics and Cognition* 10(1/2), 27–55.
- Baldi, Sergio. 2008. *Dictionnaire des emprunts arabes dans les langues de l'Afrique de l'Ouest et en swahili*. Paris: KARTHALA.
- Blench, Roger & Stuart McGill (eds.). 2012. *Advances in Minority Language Research in Nigeria*. Köln: Rüdiger Köppe Verlag.
- Beyer, Klaus. 2009. "Double negation-marking: A case of contact-induced grammaticalization in West Africa?". In: Norbert Cyffer, Erwin Ebermann & Georg Ziegelmeyer (eds.). *Negation Patterns in West African Languages and Beyond*. Amsterdam and Philadelphia: John Benjamins Publishing Company, 205-222.
- Cobbinah, Alexander. 2010. "Casamance as an area of intense language contact". *Journal of language contact* *THEMA* 3, 175–201.
- Cyffer, Norbert. 2009. "Introduction". In: Norbert Cyffer, Erwin Ebermann & Georg Ziegelmeyer (eds.). *Negation Patterns in West African Languages and Beyond*. Amsterdam and Philadelphia: John Benjamins Publishing Company, 1-6.
- Cyffer, Norbert, Erwin Ebermann & Georg Ziegelmeyer (eds.). 2009. *Negation Patterns in West African Languages and Beyond*. Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Cyffer, Norbert & Georg Ziegelmeyer (eds.). 2009. *When Languages Meet. Language Contact and Change in West Africa*. Köln: Rüdiger Köppe.
- Clements, Nick G. & Annie Rialland. 2008. "Africa as a phonological area". In: Bernd Heine & Derek Nurse (eds.). *A Linguistic Geography of Africa*. Cambridge: Cambridge University Press, 36-85.
- Dimmendaal, Gerrit. 2008. "Language Ecology and Linguistic Diversity on the African Continent". *Language and Linguistics Compass* 2, 840-858.
- Dryer, Matthew S. & Martin Haspelmath (eds.). 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology (Available online at <http://wals.info>, Accessed on 2020-07-31).
- Faraclas, Nicholas Gregory. 1989. *A grammar of Nigerian Pidgin*. Ph. D. Dissertation, University of California at Berkeley.
- Frąckiewicz Iga. 2019. *The Image of African languages structures in Nigerian Pidgin English*, Ph. D. Dissertation, University of Warsaw.
- Güldemann, Tom. 2008. "The 'Macro-Sudan belt': Towards identifying a linguistic area in northern Sub-Saharan Africa". In: Bernd Heine & Derek Nurse (eds.). *A linguistic geography of Africa*. Cambridge: Cambridge University Press, 151-185.
- Heine, Bernd & Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.

- Hollington, Andrea. 2015. *Traveling Conceptualizations. A cognitive and anthropological linguistic study of Jamaican*. Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Jaggard, Philip & Malami Buba. 2009. "Metaphorical extensions of 'eat' =>[OVERCOME] and 'drink' =>[UNDERGO] in Hausa". In: John Newman (ed.). *The linguistics of eating and drinking*. Amsterdam and Philadelphia: John Benjamins Publishing Company, 229-252.
- Jespersen, Otto. 1909-49. *Modern English Grammar on Historical Principles* (7 vols). London: Allen & Unwin.
- Krings, Matthias. 2010. "Nollywood Goes East: The Localization of Nigerian Video Films in Tanzania". In: Mahir Saul & Ralph A. Austen (eds.). *Viewing African Cinema in the Twenty-First Century: Art Films and the Nollywood Video Revolution*. Athens, OH: Ohio University Press, 74-91
- Lüpke, Friderike & Mary Chambers. 2010. "Multilingualism and language contact in West Africa: towards a holistic perspective". *Journal of Language Contact* *THEMA* 3, 1-12.
- Mazzoli, Maria. 2013. *Copulas in Nigerian Pidgin*. Ph. D. Dissertation, University of Pavia.
- Newman, Paul. 2012. "Pluractional verbs: An overview". In: Patricia Cabredo Hofherr and Brenda Laca (eds.). *Verbal Plurality and Distributivity*. Berlin: de Gruyter, 185-209.
- Polzenhagen, Frank. 2007. *Cultural Conceptualisations in West African English*. Frankfurt: Peter Lang.
- Pullum, Geoffrey K. & Rodney Huddleston. 2002. "Prepositions and preposition phrases". In: Rodney Huddleston & Geoffrey K. Pullum et al. (eds.). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press, 597-661.
- Ross, Malcolm. 1997. "Social networks and kinds of speech-community event". In: Roger Blench & Matthew Spriggs (eds.). *Archeology and Language*. London: Routledge, 209-261.
- Winkelmann, Kerstin & Gudrun Mieke. 2009. "Negation in Gur: Genetic, areal and unique features". In: Norbert Cyffer, Erwin Ebermann & Georg Ziegelmeyer (eds.). *Negation Patterns in West African Languages and Beyond*. Amsterdam and Philadelphia: John Benjamins Publishing Company, 167-204.
- Wolff, Ekkehard & Doris Löhr. 2005. "Convergence in Saharan and Chadic TAM Systems, *Afrika und Übersee* 88, 265-299.
- Wolff, Ekkehard (ed.). 2019. *The Cambridge Handbook of African Linguistics*. Cambridge: Cambridge University Press.
- Ziegelmeyer, Georg. 2009. "Negation of non-indicative mood in Hausa, Fulfulde and Kanuri". In: Norbert Cyffer, Erwin Ebermann & Georg Ziegelmeyer (eds.). *Negation Patterns in West African Languages and Beyond*, 7-20.
- Ziegelmeyer, Georg. 2015. "Areal diffusion in the Chadic-Kanuri contact zone". Paper presented at WOCAL8, Kyoto 2015 (Available online at http://idiatov.mardi.myds.me/WOCAL8_Areal_Phenomena_NSSA/ZIEGELMEYER_2015_WS_Areal_phenomena_in_NSSA_SLIDES.pdf, Accessed on 2020-07-31).
- Zima, Petr. 2009. *The Verb and Related Areal Features in West Africa*, München: LINCOM Europa.

Part I



West Africa as a linguistic area

Measuring phonological complexity in West African languages

Abstract

The discussion on the complexity of natural language is a fascinating topic that has been treated by scholars from different philosophical and theoretical perspectives. The main challenge to overcome when studying complexity is represented by its quantification: discussing complexity means dealing with objective measurements. Since languages are systems, i.e. they are made up of elements, it is possible to examine the structural complexity of a language by counting the elements present in the system. Systems (that is, languages) are in turn made of sub-systems (that is, areas), each sub-system being described by a series of features whose inventory sizes can be, for example, relatively small, relatively large, or average. This paper aims at formulating an *Index of Phonological Complexity* (IPC) based on the typological features covering the phonological area as defined in the *World Atlas of Language Structures* (WALS). After a brief discussion on these features, their recoding and subsequently their normalisation to a common scale will be argued for. Then, three indexes of phonological complexity will be proposed and applied to West African languages. Given its high degree of linguistic diversity determined by both genetic and typological variety, West Africa is an interesting ground for measuring complexity as well as a promising laboratory for further calibration and refinement of the indexes.

Keywords: phonology, complexity, West Africa

1. Complexity: what is it?

The idea that languages are somehow ‘complex’ has a relatively long history. In the 19th century, languages were seen as products of the communities that spoke them. The principle was clear: sophisticated communities used sophisticated languages and the most sophisticated communities were those represented by nations. Since the complexity of a given language was measured against abstractions such as the ‘spirit of the nation’, languages of nationless communities were deemed – here the lexical choice is strictly dependent on the context – simpler, primitive, or inferior. Races that built nations had something that nationless groups did not have: the ability to express abstractions (cf. Herder 1772; von Humboldt 1836). This (hard-to-die) idea started fading away, at least in the scientific community, by the second half of the following century, when the Academia shifted away from the fatally romantic assumptions of the past and replaced the old ideology with a new one: languages, it was said now, do not differ much from each other in terms of complexity. The over quoted passage of Charles Hockett according to which “it would seem that the total grammatical complexity of any language [...] is about the same of any

other” (Hockett 1958: 180) became emblematic of the new *Zeitgeist* permeating the view of linguistics on the differences between languages. The rise of the generative school in the 1960s neutralised the problem: the postulated existence of a universal grammar innate to *Homo sapiens* made all the observable differences between natural languages accidental and negligible. Saying that all languages are equally complex is, from a generativist point of view, true and beyond the point at the same time.

The topic of linguistic complexity, however, did not die out. Descriptions of single languages continued to be carried out and typology – also by virtue of its empirical foundation – became the right standpoint from which to observe what is going on in the realm of natural languages. More specifically, typology tells us what languages have and what they do not have, not only feature-wise but also in terms of inventories and their sizes.

The notion of linguistic complexity as an object of study in its own is newer. A significant number of publications appeared over the last decade and scholars from different theoretical and methodological frameworks have proposed a variety of approaches with the aim to address linguistic complexity as a meaningful field of enquiry (among others, Miestamo et al. 2008; Sampson et al. 2009; Baechler & Seiler 2016). These studies define complexity in a quite homogenous way, namely by operating a clear-cut distinction between *relative* and *absolute* complexity.

Relative complexity is close to the popular notion that sees complexity as mostly related to the categorization of natural languages into ‘easy languages’ vs. ‘difficult languages’, which translates into ‘easy-to-learn languages’ vs. ‘difficult-to-learn languages’. The widespread idea according to which complex languages are difficult – i.e. complexity equals difficulty – hides a more pragmatic concern: how difficult is it for a speaker of language X to learn language Y? We are all familiar with statements like

*German is easier to learn than French if your mother tongue is Dutch.
A Russian will not have so much trouble in mastering Serbian.
Oh, you are Italian? Perfect, so you understand Spanish too!*

This notion of complexity is based on the perception people have of the language they speak (source language) and the language they want to learn (target language). This kind of comparison is essentially empirical and is built upon unsurprising similarities between source and target – or, better, upon similarities between homologous systems of source and target respectively (X phonology vs. Y phonology, X morphology vs. Y morphology, and so on). Rather unsurprisingly, when (in)direct experience does not support our stereotyped understanding of complexity (=when the operation of relativizing complexity falls short of data) we get lost:

Oh boy, she speaks Mekens!¹ Tuvinian² will be no trouble at all.

Complexity can be understood from three different perspectives: cognitive, developmental, and absolute. Cognitive complexity relates to the processing costs attached to

¹ Tupian, Tuparic (South America), ISO 639-3 skf.

² Turkic, South Siberian Turkic, ISO 639-3 tyv.

linguistic structures, while developmental complexity to the way and the order of acquisition of such structures (Pallotti 2015: 117-118). These two kinds of complexities are relative to the speaker/user/learner.

Structural complexity (the kind of complexity we are interested in), on the contrary, deals with the ‘absolute’ and is defined independently of the speaker/user/learner. This notion of complexity is based on (a) the number of elements within a given system, and (b) the relations of these elements within the system. There is no correlation between ‘absolute’ complexity and ‘relative’ complexity: from an acquisitional point of view, it is perfectly normal to have structurally simple systems that are very hard to manage cognitively.

Complexity will be treated here as something that can be measured: the elements of different areas of the grammars can be ‘counted’. In phonology, such notion of absolute or grammatical complexity translates into inventory sizes. An inventory with a large number of items will be labelled [+complex], whereas an inventory with a small number of items will be labelled [-complex]. This principle is exemplified by the vowel systems of Kushi (Afroasiatic, Chadic, Nigeria, ISO 639-3 kuh) and Diyari (Pama-Nyungan, Australia, ISO 639-3 dif): the vowel system of Kushi is more complex than that of Diyari.

<table border="1"> <tr><td>i</td><td></td><td></td><td>u</td></tr> <tr><td></td><td>ɪ</td><td>ʊ</td><td></td></tr> <tr><td></td><td>e</td><td></td><td>o</td></tr> <tr><td>ə</td><td></td><td></td><td></td></tr> <tr><td></td><td>ɛ</td><td></td><td>ɔ</td></tr> <tr><td></td><td></td><td>a</td><td></td></tr> </table> <p style="text-align: center;">Kushi: [+complex]</p>	i			u		ɪ	ʊ			e		o	ə					ɛ		ɔ			a		<table border="1"> <tr><td>i</td><td></td><td>u</td></tr> <tr><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td></tr> <tr><td></td><td></td><td>a</td></tr> </table> <p style="text-align: center;">Diyari: [-complex]</p>	i		u																		a
i			u																																											
	ɪ	ʊ																																												
	e		o																																											
ə																																														
	ɛ		ɔ																																											
		a																																												
i		u																																												
		a																																												

Table 1 – Vowel systems of Kushi (Afroasiatic, Chadic / Nigeria) and Diyari (Pama-Nyungan / Australia): [±complexity]

The current debate on the complexity of languages has been focusing on the definition and metrics of complexity, challenging Hockett’s assumption on ‘total complexity’. Once the idea that a language is a mechanism whose components can be analysed in terms of numbers (later on we will see what this exactly means) is accepted, then the master question becomes: ‘how can we measure the global complexity of languages?’.

Let’s step back to the ‘total complexity’ notion stated by Hockett. The notion that sees all languages as equally complex has a strong theoretical implication. If the overall complexity of a language X is equal to the overall complexity of a language Y, then there must a compensation effect in play between different areas of the grammar. Let’s suppose, for example, that language X has a very simple phonology (e.g. 3 vowels, 17 consonant phonemes, no tones, and simple syllables such as V and CV). Then, according to the idea that all languages are equally complex, one would expect other areas of the grammar, such as morphology or syntax, to compensate for the poor phonological system. Analogously, if we took into account a language Y displaying a complex phonological system, then

we would expect other areas of the grammar to be simple or less complex. But the equality principle, reasonable (and politically-correct) it may seem, has two weak points. First of all, it is not true that all languages are equally complex: creoles, for example, are consistent in showing a ‘simpler’ grammar (Parkvall 2008). Secondly, the high complexity shown by certain languages in some areas of their grammars, e.g. verbal morphology, can’t be equalled or compensated for by complex areas in other languages. In other words, there are cases where complexity is so high that any comparison based on the equality principle does not hold: some languages are simply more complex than others.

This paper builds on previous work on linguistic complexity (Bentz et al. 2016; Maddieson 2005; 2006; 2011) and addresses the possibility to measure phonological complexity from a typological standpoint, i.e. a global quantification specifying the phonological complexity of a cluster of typological features. This paper has two main purposes: the formulation of an *Index of Phonological Complexity* (IPC), and the application of such an index across West African languages. The present study is based on the typological features defined in the *World Atlas of Language Structures* (Dryer & Haspelmath 2013), namely the chapters on phonology (Maddieson 2013a-k; Anderson 2013; Goedemans & van der Hulst 2013a-d; and Hajek 2013). In order to define an IPC, I will shortly present the phonological features considered in the WALS. Then I will argue for a numerical recoding of these features and for their subsequent normalisation to a common scale. Having completed these preliminary steps, I will propose three indexes of phonological complexity and discuss their application to West African languages.

2. West Africa: a phonological zone and a testing ground

The rationale for the choice of West Africa as the testing ground for an index of phonological complexity is due to the extreme variety of the region in terms of phonological features. While an index is by definition universal (it is possible to apply it in any context to obtain an objective measurement), not all contexts in which an index can be used are equally interesting: the most productive way to test an index of phonological complexity would be its application in a context displaying a high density of languages belonging to different language phyla. The purpose of the index (and the comparative analysis of the results obtained through its application across the languages present in a given area) does not consist in highlighting the presence of this or that phonological feature, but rather in identifying the inter-dependence of features and how their inventory sizes increase or decrease when specific features coexist in the same language.

West Africa is defined here as the region bordered by the Atlantic Ocean to the west and the south, the Sahara desert to the north, and Cameroon/Chad to the east. The region represents only a portion of the ‘Sudanic belt’ – a vast phonological zone including West Africa and extending to Lake Albert and the Ethiopian-Eritrean highlands to the south-east and east respectively – identified by Clements & Rialland (2007). Clements & Rialland’s subdivision of Africa in six phonological zones (North, East, Sudanic, Center, South, and Rift) is motivated by the observation that many phonological features specific to the African continent are geographically restricted, hence the necessity to define different zones on the basis of different clusters of phonological (i.e. segmental and prosodic) features. Almost

all the phonological features that characterise the Sudanic belt as a whole are also found in West Africa: if we accept the existence of a Sudanic belt (and certainly the evidence is compelling), then West Africa alone is an excellent representative of such phonological zone. The specificity of the West African/Sudanic region in terms of phonology is given by the presence of a) labial flaps, b) labial-velar stops, c) vowel and consonant nasality, d) ATR-based vowel harmony, and e) simple and complex tone systems.

Since West Africa is not treated here as a phonological zone *sensu stricto*, but rather as a meaningful testing ground where complexity can be measured and analysed (and, in a sense, observed in action) by virtue of a high degree of linguistic diversity, an important reason for basing a study on this region is the availability of data. Despite the fact that the scientific community is still lacking a thorough description of many African languages, the West African region, when it comes to phonological features, is fairly represented in the WALs and in print publications, thus allowing for a firm approach to phonology-based phenomena.

3. Phonological features

The WALs covers phonology by specifying 20 features:

#	WALS features: phonology	
1	1A	consonant inventories
2	2A	vowel quality inventories
3	3A	consonant-vowel ratio
4	4A	voicing in plosives and fricatives
5	5A	voicing and gaps in plosive systems
6	6A	uvular consonants
7	7A	glottalised consonants
8	8A	lateral consonants
9	9A	velar nasal
10	10A	vowel nasalisation
11	10B	nasal vowels in West Africa
12	11A	front rounded vowels
13	12A	syllabic structure
14	13A	tone
15	14A	fixed stress locations
16	15A	weight sensitive stress
17	16A	weight factors in weight-sensitive stress systems
18	17A	rhythm types
19	18A	absence of common consonants
20	19A	presence of uncommon consonants

Table 2 – WALs features for the phonological area

Each of these features displays a set of values. For a certain number of features the values are strictly related to the size of the reference inventory. This is the case of feature 1A (consonant inventories), 2A (vowel quality inventories), 12A (syllabic structure), and 13A (tone). The following tables illustrate the set of labels/values assigned to each of these features and the respective inventory size.

1A – WALS value	Number of consonants
Small	6-14
Moderately small	15-18
Average	22 (±3)
Moderately large	26-33
Large	34≤

Table 3 – Consonant: inventory types

2A – WALS value	Number of vowels
Small	2-4
Medium	5-6
Large	7-14

Table 4 – Vowels: inventory types

12A – WALS value	Syllable types
Simple	(C)V, CV
Moderately simple	CVC, CCV, CCVC
Complex	(C)(C)(C)V(C)(C)(C)(C)

Table 5 – Syllabic structures

13A – WALS value	Tone
No tone	–
Simple tone system	two-way contrast
Complex tone system	3≤

Table 6 – Tones

Rather than being tied to the size of an inventory, some features are defined on the basis of the presence (or absence) of a certain phonological element and in the way this element surfaces in the language. See for example feature 7A (glottalised consonants):

7A – WALS value
No glottalised consonants
Ejectives only
Implosives only

Glottalised resonants only
Ejectives and implosives
Ejectives and glottalised resonants
Implosives and glottalised resonants
Ejectives, implosives, and glottalised resonants

Table 7 – Glottalised consonants

4. Measuring phonological complexity in a typological perspective

In order to obtain an index of phonological complexity, i.e. a value measuring the complexity of a cluster of features, it is necessary (1) to recode the values used in the WALS by converting them into a numerical scale, and (2) to normalise these new values to a common scale.

4.1. Values: recoding

The WALS categorises consonant, vowel, syllabic, and tone inventories according to their sizes and/or complexity. Consonant and vowel inventories are defined in terms of size, whereas syllabic structures and number of tones are referred to in terms of $[\pm$ complexity]. We can operate a label normalisation and think the four systems in terms of structural complexity, i.e. we could range the size types of the four inventories on a scale that goes from [+simple, –complex] to [–simple, +complex], where a system’s complexity and the number of its elements are directly proportional. Since the aim is to attain a measure of the phonological complexity in a given language with the ultimate goal to compare values, the labels will be recoded in terms of numeric values. In the tables below, numeric values have been assigned to the WALS inventory types – the lower the rank in terms of complexity, the lower the assigned value. The five types of consonant inventories have been arranged on a scale from 1 [+simple, –complex] to 5 [–simple, +complex].

WALS value	Recoded value
small	1
mod. small	2
average	3
mod. large	4
large	5

Table 8 – WALS values for consonant inventories coded on a [1, 5] scale

Values to vowel, syllabic, and tonal inventories have been assigned in the same way. These phonological systems display three major groups the inventories can be categorised into, hence the values have been assigned on a [1, 3] scale:

Vowels	Value	Syllables	Value	Tones	Value
small	1	simple	1	∅	1
medium	2	mod. simple	2	simple	2
large	3	complex	3	complex	3

Table 9 – WALs values for vowel/syllable/tone inventories coded on a [1, 3] scale

Other features whose values have been rearranged on a [1, 3] scale are 4A (voicing in plosive and fricatives), 5A (voicing and gaps in plosive systems), and 6A (uvular consonants). For these features it is not possible to assign a different score to each of the intermediate values, as their degree of complexity is the same. Consider, for example, the case of feature 4A. The WALs defines 4 values, but not 4 different degrees of complexity. The absence of voicing contrast represents the minimum degree of complexity, the presence of contrast in plosive and fricative the maximum degree, but the plosives-only contrast and the fricatives-only contrast are undistinguishable in terms of complexity, hence they have been quantified with the same score. A four-value set has been coded to a three-value range.

4A – WALs values	WALs value [1, 3]
no voicing contrast	1
voicing contrast in plosive alone	2
voicing contrast in fricatives alone	2
voicing contrast in both plosives and fricatives	3

Table 10 – WALs values for plosive contrast coded on a [1, 3] scale

Similarly, the five-value feature 8A has been coded on a three-degree scale:

8A – WALs values	WALs value [1, 3]
no laterals	1
no /l/, but lateral obstruents	2
laterals, but no /l/, no obstruent laterals	2
/l/, no obstruent laterals	2
/l/ and lateral obstruent	3

Table 11 – WALs values for laterals [1, 3] scale

The seven-value feature 19A (presence of uncommon consonants) has been transformed into a binary system ([0, 1]):

19A – WALs values	WALs value [0, 1]
None	0
Pharyngeals and “th”	1
Pharyngeals	1

None	1
Labial-velars	1
Clicks, pharyngeals, and “th”	1
Clicks	1
“Th” sounds	1

Table 12 – WALS values for the presence of uncommon consonants on a binary [0, 1] scale

#		WALS feature	Scale	Type	Transformation
1	1A	consonant inventories	[1, 5]	ordinal	[0, 1]
2	2A	vowel quality inventories	[1, 3]	ordinal	[0, 1]
3	3A	consonant-vowel ratio	[1, 5]	ordinal	[0, 1]
4	4A	voicing in plosives and fricatives	–	mixed	[0, 1] + reordered
5	5A	voicing and gaps in plosive systems	–	mixed	[0, 1] + reordered
6	6A	uvular consonants	[1, 3]	mixed	[0, 1]
7	7A	glottalised consonants	–	mixed	[0, 1] binary
8	8A	lateral consonants	–	mixed	[0, 3] reordered
9	9A	velar nasal	–		[0, 1] binary
10	10A	vowel nasalisation	[1, 2]	binary	[0, 1] binary
11	10B	nasal vowels in West Africa	–	mixed	[0, 1] reordered
12	11A	front rounded vowels	–	mixed	[0, 1] reordered
13	12A	syllabic structure	[1, 3]	ordinal	[0, 1]
14	13A	tone	[1, 3]	ordinal	[0, 1]
15	14A	fixed stress locations	–	mixed	[0, 1] binary
16	15A	weight sensitive stress	–	mixed	[0, 1] binary
17	16A	weight factors in weight-sensitive stress systems	–	mixed	[0, 1] binary
18	17A	rhythm types	–	mixed	[0, 1] binary
19	18A	absence of common consonants	–	mixed	[0, 1] binary
20	19A	presence of uncommon consonants	[0, 1]	binary	[0, 1] binary

Table 13 – Transformations

4.2. Values: normalisation

In order to be comparable, the value ranges of the scales obtained through recoding ([1, 5], [1, 3] and [0, 1]) need to be normalised, i.e. to be transformed and arranged on a common scale. The three scales will be normalised to a [0, 1] interval. The reason behind the choice of the [0, 1] interval is that it will allow us to express the Index of Phonological Complexity as a fraction of 1, i.e. 0 followed by n digits. Moreover, the [0, 1] interval is consistent with the previous literature on linguistic complexity (Bentz et al. 2016) or diversity (Harmon & Loh 2010) indexes.

$$(1) \quad \text{Normalised}(x_i) = \frac{x_i - X_{min}}{X_{max} - X_{min}}$$

Where:

X_{min} = the minimum value for variable X

X_{max} = the maximum value for variable X

x_i = the raw data to be normalised

As a result, the normalised values are as follows:

Recoded WALS values			New values
[1, 5]	[1, 3]	[0, 1]	[0, 1]
1	1	0	0
2			0.25
3	2		0.5
4			0.75
5	3	1	1.0

Table 14 – Normalised values: correspondences

5. Index of phonological complexity

An IPC can be obtained by calculating the feature value average per language:

$$(2) \quad \text{IPC}_{WALS(20)} = \frac{\sum_{i=1}^n f_i}{n}$$

Where:

f_i = value of feature i

n = number of features available per language

The formula takes into account only the features for which the WALS assigns a value. Table A in the Appendix lists the West African languages present in the WALS according to their IPC (in descending order, where the last column indicates the number of features available).

5.1. Adjusting the index

The IPC based on the 20 phonological features specified in the WALS poses some questions. First of all, as shown in the table A of the Appendix, for many features there are no values. Then, it is reasonable to say that a complete picture of the values of phonological features is – at least with the quality and amount of data available at this moment – very difficult to attain. These two immediate observations lead to the following question: how many features do we have to take into account in order to calculate a reliable IPC? I will discuss three main options: (a) a comprehensive index, (b) a quantitative-qualitative index

based on the current availability of data, and (c) a ‘quick’ index based on a very limited number of features.

Comprehensive index based on 20 features:

The most complete index we can envisage is an index that takes into consideration all the phonological features. This index, the $IPC_{WALS(20)}$, is calculated using the formula indicated below.

$$(3) \quad IPC_{WALS(20)} = \frac{\sum_{i=1}^n f_i}{n} \quad (\text{where } 1 \leq n \leq 20)$$

The $IPC_{WALS(20)}$ is certainly comprehensive, but its extent seems to collide with the paucity of data in some phonological sub-areas, e.g. features such as 14A, 15A, 16A and 17A are rarely described and no value can be assigned to them. Moreover, this index considers both quantitative and qualitative features, whereas only strictly quantitative features are indicative of grammatical complexity.

Quantitative-qualitative index based on 13 features:

The $IPC_{WALS(13)}$ mirrors the data available in the WALS. This index leaves aside the features that are described only sporadically, considering instead the feature values that are statistically consistent in terms of presence in the WALS. Out of 20 features, we observe that 13 are those that make it to the WALS: 1A, 2A, 3A, 4A, 5A, 6A, 7A, 8A, 11A, 12A, 13A, 18A, 19A. While the $IPC_{WALS(13)}$ considers a relatively high number of features, it stills mixes quantitative and qualitative features.

$$(4) \quad IPC_{WALS(13)} = \frac{\sum_{i=1}^n f_i}{n} \quad (\text{where } n = 13)$$

Quantitative index based on 4 features:

The last index is perhaps the most intuitive. It cuts the issue down to size by considering only 4 features: consonant inventory (1A), vowel inventory (2A), syllabic structure (12A), and tonal system (13A).

$$(5) \quad IPC_{WALS(4)} = \frac{\sum_{i=1}^n f_i}{n} \quad (\text{where } n = 4)$$

The $IPC_{WALS(4)}$ is strictly quantitative. It takes into account only core features whose values are highly documented. This index does not consider dependent features and values (for example, consonant-vowel ratio) and bypasses the paucity of data in certain phonological sub-areas (e.g. prosody-related features). Nevertheless, the $IPC_{WALS(4)}$ relies heavily on the experts’ judgement: the incorrect description of certain features (e.g. failure to recognise vowel harmony or tonal systems) will irremediably distort the reading of the complexity measure.

5.2. Representativeness

In the previous section I have described three possible indexes. Each of these three metric tools presents advantages and disadvantages. In order to operate a choice we are

forced to look for a master criterion, which I claim to be representativeness. Even if we could assign a value to all the 20 features listed in the WALs, the obtained phonological picture is deemed to be a coarse-grain approximation. Languages can be categorised or grouped together according to the values assigned to certain features, indeed the main function of these features is dividing languages into types and not describing areas of grammar in detail. Some features are more representative than others: we could say, for example, that knowing the size of a consonant inventory is more relevant than assessing the presence or absence of velar nasals. Hence, we could tie the notion of representativeness to quantifiable inventories (consonants and vowels) and structures (syllables and tones), leaving asides binary features (velars nasals, front rounded vowels, etc.). In other words, representative features are independent, i.e. they express inventories and structures and not parts of them. Therefore, among the three indexes proposed here, the $IPC_{WALS(4)}$ (which, at this point, could be renamed IPC_4) is the most representative in that it considers only the most important (i.e. basic and quantifiable) features.

6. Phonological complexity in West African languages: some observations

In this section I will address some issues of methodological and theoretical order arising from the application of the ICP. The first observation illustrates a general principle, labelled here “Complexity Equilibrium”, while the others relate to the specific case of West African languages.

6.1. The Complexity Equilibrium

The typological data available in the WALs tell us that no language displays maximum or minimum degrees of complexity.

Consonants	Vowels	Syllables	Tones	<i>n</i> of languages
large (34+)	large (7-14)	complex	complex	0
mod. small. (15-18)	large (7-14)	complex	complex	0
small (6-14)	large (7-14)	complex	complex	0
small (6-14)	small (2-4)	simple	simple	0

Table 15 – Maximum/minimum complexity (across languages present in WALs for which data are available for the 4 phonological factors)

While it is perfectly normal to find systems (or factors) with either a very small or a very large number of elements, when it comes to the configuration of systems (i.e. all the systems relevant, say, to phonology) the high complexity of one factor must be balanced by the low complexity of another factor. In this sense, a macro-system (e.g. phonology, morphology) can be neither ‘overloaded’ nor ‘under loaded’.

6.2. Compensation effect

Maddieson, in his statistical analysis on the relationship between syllable structures, segment inventories and tone contrasts observes that no significant ‘compensation effect’ can be found between an increase of complexity in one sector of phonology and a decrease of complexity in another sector, the only exception being the relationship between tonal systems and syllabic structures (Maddieson 2007: 93ff.). He also states that “increasing complexity of tone system is positively associated with increasing size of both consonant and vowel inventories” (Maddieson 2007: 102).

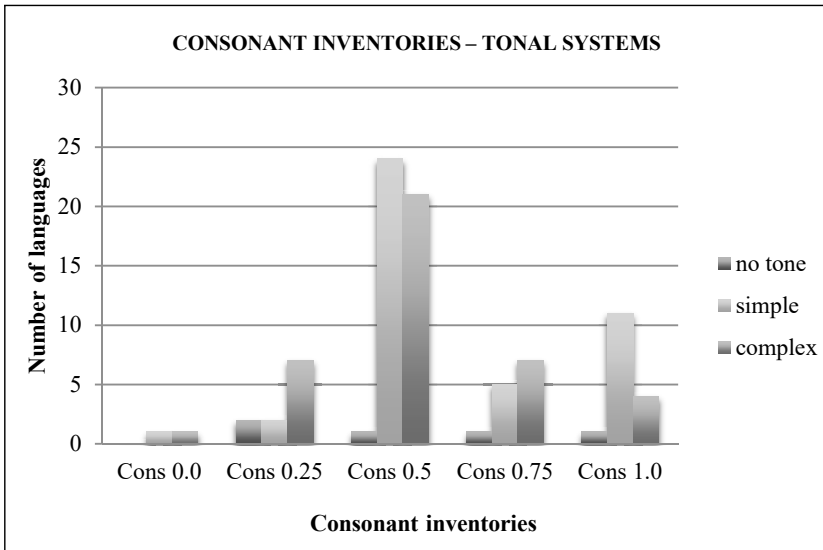


Figure 1 – Consonant inventories and tonal systems

In this sense, West African languages seem to diverge from the general pattern identified by Maddieson. Out of 98 West African languages (WALS plus other sources), 46 languages present average consonant inventories. Of these, 45 have either a simple tonal system (two-way contrast) or a complex tonal system (number of tones ≥ 3). The total number of languages with a large consonant inventory is 16, 11 of which display a simple tonal system and 4 a complex tonal system.

6.3. Availability of data

The tables displayed in the Appendix consider exclusively the languages for which the WALS reports the four core features, disregarding all those languages without a description of consonant inventory, vowel inventory, syllable structure, and tonal system. The West African languages listed in the WALS that satisfy this criterion amount to 56, which is a relatively small number if compared to the total number of West African languages represented in the *Atlas*. Among the languages for which the WALS does not provide any phonological data there is a high number of under-described and undescribed languages, but also vehicular languages such as Kanuri or Pulaar. Although the IPC_{WALS(4)} requires

a limited number of features, these are not always readily available in the WALs. Fortunately, core features can be easily assessed by consulting (when existing!) grammatical sketches, articles, fieldwork reports, and conference papers.³ One of the major features of the WALs is that all languages are geolocalised (although not necessarily in a precise manner), hence it is possible to sort them by feature or combination of features and have them represented in a map according to the values of the selected features. Considering both the WALs and a certain number of external sources, it seems that there is a general lack of data concerning the prosodic aspects of languages. Table A in the Appendix shows how the features “fixed stress locations” (14A), “weight sensitive stress” (15A), “weight factors in weight-sensitive stress systems” (16A), and “rhythm types” (17A) are rarely described. Other features that are left undescribed in the WALs do not entail paucity of data: the absence of description for certain features has more to do with the way the data has been collected and/or with the timing of its insertion in the database than with the nature of the data or the source itself. If we have at our disposal the values for features such as consonant and vowel inventories, then it is reasonable to think that the lack of values regarding the presence or the absence of velar nasals (9A), vowel nasalisation (10A), and nasal vowels (10B) is merely accidental (in other words, this is a case where the lack of information can be overcome by accessing the primary source).

6.4. West Africa as a phonological area

The purpose of this paragraph is to summarise the areal aspects from the perspective of phonological complexity as resulting from the application of the IPC (see Appendix). Table A shows the results of the $IPC_{WALS(20)}$ across the 56 West African languages listed in the WALs for which at least the core features are valued (that is, languages with no indication of any of the core features have been ignored). The languages score from a maximum of 0.700 (Doyayo) to a minimum of 0.269 (Ifik). Using different IPC, languages not only score differently but are also grouped differently (Appendix, table B). The application of the $IPC_{WALS(20)}$ results in 31 positions: languages scoring 0.700 (1 member: Doyayo), languages scoring 0.615 (4 members: Angas, Gwari, Kpelle, Tera), and so on. The $IPC_{WALS(13)}$ results in 20 positions, while the $IPC_{WALS(4)}$ results in just 12. For obvious reasons there is an observable general correspondence between the $IPC_{WALS(20)}$ and the $IPC_{WALS(13)}$. Nevertheless, if we take into account the $IPC_{WALS(20)}$ and the $IPC_{WALS(4)}$ we will observe, at least for certain languages, a certain degree of idiosyncrasy. Fe’fe (Volta-Congo), for example, scores 0.346 (29th/31) with the $IPC_{WALS(20)}$ (max. value 0.700, min. value 0.269) and 0.688 (7th/12) with the $IPC_{WALS(4)}$ (max. value 0.938, min. value 0.313). Is the phonology of Fe’fe non-complex, as the $IPC_{WALS(20)}$ seems to suggest, or is this language relatively complex, as the calculation with $IPC_{WALS(4)}$ implies? Again, the entire issue goes down to the representativeness of features. The reason for such a low score with the $IPC_{WALS(20)}$ resides in the fact that the values of a few non-core features lower the global score in a significant way. This does not mean that phonological complexity cannot be properly measured, but rather that more than one index is possible and that different indexes are based on different phonological features (even if, as argued before, an index based on core features appears to be the best

³ Bibliography of data-related literature not present in the WALs is listed at the end of the article.

candidate in order to attain a global, objective quantification of complexity). West African languages are consistent in their syllabic structures, which generally are moderately simple, while showing a less homogeneous pattern in consonant and vowel inventories and in tonal systems.

7. Conclusions

In this paper I have argued for the existence of a tool – in *Index of Phonological Complexity* – capable of measuring the phonological complexity of any given language. I have used the index, based on the phonological features listed in the WALS, to calculate the phonological complexity in a certain number of West African languages. The main purpose of the IPC is to calculate phonological complexity cross-linguistically by taking into account the data provided by the *highest number of languages possible*. As we have seen, the WALS, while offering a solid typological framework to build the metrics of the IPC, does not include all the data concerning phonologies of individual languages that are available in the literature. Hence, to shift from a macro-scale scenario to a micro-scale scenario, and then to a detailed representation of the latter, more data (both from existing material and from the field) are needed. A general grasp on complexity is, at least within certain limits, quite intuitive in the context of West Africa, but specific – regional and areal – dynamics require an up-to-date and detailed representation. IPC values, in order to say something meaningful, need to be statistically sound and anchored to phenomena and conditions that can play a role in the increase or decrease of complexity: geographical setting, language contact dynamics, and population size, just to name a few. In this sense, West Africa is a particularly challenging area. The lack basic phonological descriptions for a high number of languages hinders a meaningful representation of complexity, especially as far as extra-linguistic factors are concerned.

References

- Anderson, Gregory D.S. 2013. “The Velar Nasal”. In: M.S. Dryer, M. Haspelmath (eds.). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/9>, Accessed on 2018-08-24.
- Baechler, Raffaëla & Guido Siler. 2016. *Complexity, Isolation, and Variation*. Berlin & Boston: de Gruyter.
- Bentz, Christian, Tatyana Ruzsics, Alexander Kopenig, & Tanja Samardžić. 2016. “A Comparison Between Morphological Complexity Measures: Typological Data vs. Language Corpora”. In: Dominique Brunato, Felice Dell’Orletta, Giulia Venturi, Thomas François, and Philippe Blache (eds.). *Proceedings of the Workshop on Computational Linguistics for Linguistic Complexity*. Osaka: The COLING 2016 Organizing Committee, 143-153.
- Clements, Nickerson G. & Annie Rialland. 2007. “Africa as a phonological area”. In: Bernd Heine, Derek Nurse (eds.), *A Linguistic Geography of Africa*. Cambridge: Cambridge University Press, 36-85.
- Dryer, Matthew S. & Martin Haspelmath (eds.). 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/3>, Accessed on 2018-09-18.
- Goedemans, Rob & Harry van der Hulst. 2013a. “Fixed Stress Locations”. In: Matthew S. Dryer & Martin Haspelmath (eds.). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/14>, Accessed on 2018-08-24.

- Goedemans, Rob and Harry van der Hulst. 2013b. "Weight-sensitive stress". In: Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/15>, Accessed on 2018-08-24.
- Goedemans, Rob and Harry van der Hulst. 2013c. "Weight factors in weight-sensitive stress systems". In: Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/16>, Accessed on 2018-08-24.
- Goedemans, Rob and Harry van der Hulst. 2013d. "Rhythm types. Fixed Stress Locations". In: Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/17>, Accessed on 2018-08-24.
- Hajek, John. 2013. "Velar nasalization". In: Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/10>, Accessed on 2018-08-24.
- Harmon, David & Jonathan Loh. 2010. "The Index of Linguistic Diversity: A New Quantitative Measure of Trends in the Status of the World's Languages". *Language Documentation & Conservation* 4, 97-151.
- Herder, Johannes Gottfried. 1772. *Abhandlung über den Ursprung der Sprache: welche den von der Königl. Academie der Wissenschaften für das Jahr 1770 gesezten Preis erhalten hat*. Berlin: Voss.
- Hockett, Charles Francis. 1958. *A Course in Modern Linguistics*. New York: Macmillan.
- von Humboldt, Wilhelm. 1836. *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts*. Berlin: Ferdinand Dümmler.
- Maddieson, Ian. 2005. "Issues of phonological complexity: Statistical analysis of the relationship between syllable structures, segment inventories and tone contrasts". In: *UC Berkeley Phonology Lab Annual Report*. Berkeley: University of California, 269-268
- Maddieson, Ian. 2006. "Correlating phonological complexity: Data and validation". *Linguistic Typology* 10(1), 106-123
- Maddieson, Ian. 2007. "Issues of Phonological Complexity Statistical Analysis of the Relationship between Syllable Structures, Segment Inventories, and Tone Contrasts". In: Maria-Josep Solé, Patrice Speeter Beddor, and Manjari Ohala (eds.), *Experimental Approaches to Phonology*. Oxford: Oxford University Press, 93-103.
- Maddieson, Ian. 2011. "Phonological complexity in linguistic patterning". In: *Proceedings of the 17th International Congress of Phonetic Sciences (Hong Kong, August 17-21)*: 28-34.
- Maddieson, Ian. 2013a. "Consonant Inventories". In: Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/1>, Accessed on 2018-08-24.
- Maddieson, Ian. 2013b. "Vowel quality inventories". In: Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/2>, Accessed on 2018-08-24.
- Maddieson, Ian. 2013c. "Consonant-vowel ratio". In: Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/3>, Accessed on 2018-08-24.
- Maddieson, Ian. 2013d. "Voicing in plosives and fricatives". In: Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/4>, Accessed on 2018-08-24.
- Maddieson, Ian. 2013e. "Voicing in gaps and plosive systems". In: Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/5>, Accessed on 2018-08-24.
- Maddieson, Ian. 2013f. "Uvular consonants". In: Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/6>, Accessed on 2018-08-24.
- Maddieson, Ian. 2013g. "Glottalized consonants". In: Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/7>, Accessed on 2018-08-24.

- Maddieson, Ian. 2013h. "Lateral consonants". In: Matthew S. Dryer & Martin Haspelmath (eds.). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/8>, Accessed on 2018-08-24.
- Maddieson, Ian. 2013i. "Front rounded vowels". In: Matthew S. Dryer & Martin Haspelmath (eds.). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/11>, Accessed on 2018-08-24.
- Maddieson, Ian. 2013j. "Syllable structure". In: Matthew S. Dryer & Martin Haspelmath (eds.). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/12>, Accessed on 2018-08-24.
- Maddieson, Ian. 2013k. "Tone". In: Matthew S. Dryer & Martin Haspelmath (eds.). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at <http://wals.info/chapter/13>, Accessed on 2018-08-24.
- Miestamo, Matti, Kaius Sinnemäki & Fred Karlsson. 2008. *Language Complexity. Typology, contact, change*. Amsterdam & Philadelphia: John Benjamins.
- Pallotti, Gabriele. 2015. "A simple view of linguistic complexity". *Second Language Research* 31(1), 117-134.
- Parkvall, Mikael. 2008. "The simplicity of creoles in a cross-linguistic perspective". In: Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds.). *Language Complexity. Typology, contact, change*. Amsterdam & Philadelphia: John Benjamins, 265-286.
- Sampson, Geoffrey, David Gil & Peter Trudgill. 2009. *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press.

Bibliography of data-related literature not present in the WALS

- Alio, Khalil & Herrmann Jungrathmayr. 1989. *Lexique bidiya*. Frankfurt am Main: Vittorio Klostermann.
- Andreas, Heike, Rudolf Leger & Ulrike Zoch. 2009. "The Nyam language: First steps toward a grammatical description". In: E. Rothmaler (ed.). *Topics in Chadic Linguistics V. Comparative and Descriptive Studies*. Köln: Rüdger Köppe Verlag, 23-36.
- Baerman, Matthew, Dunstan Brown & Greville G. Corbett. 2009. "Morphological Complexity: a typological perspective". In: *European Science Foundation Exploratory Workshop on Words in Action*, 2009-10-12. Istituto di Linguistica Computazionale.
- Baerman, Matthew, Dunstan Brown & Greville G. Corbett. 2017. *Morphological Complexity*. Cambridge: Cambridge University Press.
- Barreteau, Daniel. 1988. *Description du mofu-gudur. Langue de la famille tchadique parlée au Cameroun*. Paris: ORSTOM.
- Barreteau, Daniel & Yves Le Bleis. 1990. *Lexique mafa. Langue de la famille tchadique parlée au Cameroun*. Paris: Geuthner-ORSTOM.
- Batic, Gian Claudio. 2014. *A Grammatical Sketch of Bure*. Köln: Rüdiger Köppe Verlag.
- Blench, Roger. 1998, 2004. *Dagbani-English Dictionary* (ms). Available online at <http://www.rogerblench.info/Language/Niger-Congo/Gur/Dagbani%20dictionary%20CD.pdf>, Accessed on 29.07.2018.
- Blench, Roger. 2005. *The Kayigang [=Keiga, Deiga] Language of the Nuba Hills, Sudan*. Kay Williamson Educational Foundation (UK). Available online at <http://www.rogerblench.info/Language/Nilo-Saharan/Kadu/Kayigang%20wordlist.pdf>, Accessed on 15.12.2017.
- Blench, Roger. 2007. The Dyarim language of central Nigeria and its affinities. In: Henry Tourneux (ed.). *Topics in Chadic Linguistics IV. Comparative and Descriptive Studies*. Köln: Rüdiger Köppe Verlag, 41-60.
- Blench, Roger. 2009. *Bura phonology and some suggestions concerning the orthography*. Kay Williamson Educational Foundation (UK). Available online at <http://www.rogerblench.info/Language/Afroasiatic/Chadic/Central/Bura/Bura%20phonology.pdf>, Accessed on 16.12.2017.
- Blench, Roger. 2013a. *The Nupoid Languages of West-Central Nigeria: Overview and Comparative Wordlist*. Kay Williamson Educational Foundation (UK). Available online at <http://www.rogerblench.info/Language/Niger-Congo/VN/Nupoid/Nupoid%20Overview%202013.pdf>, Accessed on 15.12.2017.
- Blench, Roger. 2013b. *Ywom wordlist*. Kay Williamson Educational Foundation (UK). Available online at <http://www.rogerblench.info/Language/Afroasiatic/Chadic/West/Ywom/Ywom%20wordlist.pdf>, Accessed on 16.12.2017.

- Blench, Roger & Antony Damsai. 2009. "An introduction to Kirya-Konzəl". In: Eva Rothmaler (ed.). *Topics in Chadic Linguistics V. Comparative and Descriptive Studies*. Köln: Rüdiger Köppe Verlag, 75-85.
- Brindle, Jonathan (2017). *A Dictionary and Grammatical Outline of Chakali*. Berlin: Language Science Press.
- Frajzyngier, Zygmunt. 1989. *A Grammar of Pero*. Berlin: Dietrich Verlag Berlin.
- Friesen, Dianne. 2017. *A Grammar of Moloko*. Berlin: Language Science Press.
- Guest, Elizabeth. 1997. *Heiban Grammar*, ms.
- Jungraithmayr, Hermann. 1991. *A Dictionary of the Tangale Language*. Berlin: Dietrich Reimer Verlag.
- Kaki, Aaid Komi, Hassan Ali & Ahmad Ibrahim. 1998. *Phonology write-up of Hadara (Lote)*, ms.
- Kastenholz, Raimund. 2017. *La Langue Bolgo du Guéra (Tchad): Notes de recherche et matériel lexical*. [Arbeitspapiere des Instituts für Ethnologie und Afrikastudien der Johannes Gutenberg-Universität Mainz, 172.] Mainz: Institut für Ethnologie und Afrikastudien, Johannes Gutenberg-Universität.
- Kidda, Mairo Elinor. 1985. *Tangale Phonology: A Descriptive Analysis (Nigeria)*. Doctoral dissertation, University of Illinois at Urbana-Champaign.
- Leger, Rudolf. 2014. "Some observations on typological features in languages of the Bole-Tangale group". In: Anne Storch, Johannes Harnischfeger, & Rudolf Leger (eds.). *Fading delimitations. Multilingual settlements in a convergence area – Case studies from Nigeria*. Köln: Rüdiger Köppe Verlag, 229-262.
- Morton, Deborah. 2012. "[ATR] Harmony in an Eleven Vowel Language: The Case of Anii". In: Michael R. Marlo, Nikki B. Adams, Christopher R. Green, Michelle Morrison, & Tristan M. Purvis (eds.). *Selected Proceedings of the 42nd Annual Conference on African Linguistics*. Somerville, MA: Cascadilla Proceedings Project, 70-78.
- Roberts, James S. 2009. "Palatalization and labialization in Mawa (eastern Chadic)". In: Eva Rothmaler (ed.). *Topics in Chadic Linguistics V. Comparative and Descriptive Studies*. Köln: Rüdiger Köppe Verlag, 129-140.
- Sampson, Geoffrey. 2009. "A linguistic axiom challenged". In: Geoffrey Sampson, David Gil & Peter Trudgill (eds.). *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press, 1-18.
- Schadeberg, Thilo C. 1985. *A Small Sketch of Ewe* [Afrikanistische Arbeitspapiere: Sondernummer, 1985]. Köln: Institut für Afrikanistik.
- Schaefer, Ronald P. & Francis O. Egbokhare. 2007. *A Dictionary of Emai. An Edoid Language of Nigeria*. Köln: Rüdiger Köppe Verlag.
- Schiattarella, Valentina. 2017. *Berber Texts from Siwa (Egypt)*. Köln: Rüdiger Köppe Verlag.
- Zima, Petr. 1994. *Lexique dendi (Songhay) avec un index français-dendi*. Köln: Rüdiger Köppe Verlag.

Appendix

LANGUAGE	ISO	FAMILY	1A	2A	3A	4A	5A	6A	7A	8A	9A	10A	10B	11A	12A	13A	14A	15A	16A	17A	18A	19A	ICP	n.	
Doyayo	dow	Benue-Congo	0.50	1.0	0.50	1.0	1.0	0.0	1.0	0.5	–	1.0	1.00	0.0	1.0	1.0	–	–	–	–	0.0	1.0	0.700	15	
Angas	anc	Chadic	0.75	1.0	0.75	1.0	1.0	0.0	1.0	0.5	–	–	–	0.0	1.0	1.0	–	–	–	–	0.0	0.0	0.615	13	
Gwari	gbr	Benue-Congo	0.50	0.5	0.50	1.0	1.0	0.0	1.0	0.5	–	–	–	0.0	1.0	1.0	–	–	–	–	0.0	1.0	0.615	13	
Kpelle	xpe	Mande	0.50	1.0	0.50	1.0	1.0	0.0	1.0	0.5	–	–	–	0.0	0.5	1.0	–	–	–	–	0.0	1.0	0.615	13	
Tera	ttr	Chadic	1.00	0.5	1.00	1.0	1.0	0.0	1.0	1.0	–	–	–	0.0	0.5	1.0	–	–	–	–	0.0	0.0	0.615	13	
Kohumono	bes	Volta-Congo	0.75	1.0	0.50	1.0	1.0	0.0	1.0	0.0	–	–	–	0.0	–	1.0	–	–	–	–	0.0	1.0	0.604	13	
Aizi	ahp	Volta-Congo	0.50	1.0	0.25	1.0	1.0	0.0	1.0	0.5	–	–	–	0.0	1.0	–	–	–	–	–	0.0	1.0	0.604	12	
Tarok	yer	Volta-Congo	0.75	0.5	0.50	1.0	1.0	0.0	1.0	0.5	–	–	–	0.0	0.5	1.0	–	–	–	–	0.0	1.0	0.596	13	
Bété	bev	Volta-Congo	0.50	1.0	0.00	1.0	1.0	0.0	1.0	0.5	–	–	–	0.0	0.5	1.0	–	–	–	–	0.0	1.0	0.577	13	
Dan	dnj	Mande	0.50	1.0	0.00	1.0	1.0	0.0	1.0	0.5	–	–	–	0.0	0.5	1.0	–	–	–	–	0.0	1.0	0.577	13	
Mumuye	mzm	Volta-Congo	0.50	0.5	0.50	1.0	1.0	0.0	1.0	0.5	–	–	–	0.0	0.5	1.0	–	–	–	–	0.0	1.0	0.577	13	
Mambila	mcu	Benue-Congo	0.50	1.0	0.25	0.5	1.0	0.0	1.0	–	–	–	–	0.0	0.5	1.0	–	–	–	–	0.0	1.0	0.563	13	
Gã	gaa	Volta-Congo	0.75	1.0	0.50	1.0	1.0	0.0	0.0	0.5	–	–	–	0.0	0.5	1.0	–	–	–	–	0.0	1.0	0.558	13	
Ogbia	ogb	Volta-Congo	0.50	1.0	0.25	1.0	1.0	0.0	1.0	0.5	–	–	–	0.0	0.5	0.5	–	–	–	–	0.0	1.0	0.558	13	
Konyagi	cou	Atlantic-Congo	1.00	–	–	1.0	1.0	0.0	1.0	0.5	–	–	–	–	–	0.5	–	–	–	–	0.0	0.0	0.556	9	
Ewe	ewe	Volta-Congo	0.75	1.0	0.50	1.0	1.0	0.0	0.0	0.5	1.0	–	–	0.0	0.5	0.5	–	–	–	–	0.0	1.0	0.554	14	
Bagirmi	bmi	Central Sudanic	0.75	1.0	0.50	0.5	1.0	0.0	1.0	0.5	1.0	1.0	–	0.0	0.5	1.0	1.0	0.0	0.0	0.0	–	0.0	0.0	0.542	18
Birom	bom	Volta-Congo	0.50	1.0	0.50	1.0	1.0	0.0	0.0	0.5	–	–	–	0.0	0.5	1.0	–	–	–	–	0.0	1.0	0.538	13	
Ngizim	ngi	Chadic	1.00	0.5	1.00	1.0	1.0	0.0	1.0	1.0	0.0	–	–	0.0	0.5	0.5	–	–	–	–	0.0	0.0	0.536	14	
Ewondo	ewo	Volta-Congo	0.75	1.0	0.50	1.0	0.5	0.0	0.0	0.5	–	–	–	0.0	0.5	1.0	–	–	–	–	0.0	1.0	0.519	13	
Ndut	ndv	Atlantic-Congo	0.75	1.0	0.50	0.5	1.0	0.0	1.0	0.5	–	–	–	0.0	0.5	1.0	–	–	–	–	0.0	0.0	0.519	13	
Kpan	kpk	Benue-Congo	0.50	0.5	0.75	1.0	1.0	0.0	0.0	0.0	–	–	–	0.0	1.0	1.0	–	–	–	–	0.0	1.0	0.519	13	
Tampulma	tpm	Volta-Congo	0.50	1.0	0.25	1.0	1.0	0.0	0.0	0.5	–	–	–	0.0	1.0	0.5	–	–	–	–	0.0	1.0	0.519	13	

LANGUAGE	ISO	FAMILY	1A	2A	3A	4A	5A	6A	7A	8A	9A	10A	10B	11A	12A	13A	14A	15A	16A	17A	18A	19A	ICP	n.
Isoko	iso	Benue-Congo	0.75	1.0	0.50	1.0	1.0	0.0	0.0	0.5	-	-	-	-	0.0	0.5	0.5	-	-	-	0.0	1.0	0.519	13
Kotoko	ktk	Chadic	0.75	1.0	0.50	0.5	1.0	0.0	1.0	0.5	-	-	0.00	0.0	1.0	0.5	-	-	-	-	0.0	1.0	0.517	15
Igbo	ibo	Benue-Congo	0.75	1.0	0.50	1.0	1.0	0.0	0.0	0.5	-	-	-	0.0	0.5	0.5	-	-	-	-	0.0	1.0	0.517	15
Grebo	grj	Volta-Congo	0.50	1.0	0.25	0.5	1.0	0.0	0.0	0.5	-	-	-	0.0	0.0	1.0	0.0	1.0	1.0	-	0.0	1.0	0.513	19
Bariba	bba	Volta-Congo	0.25	1.0	0.25	1.0	1.0	0.0	0.0	0.5	-	-	-	0.0	-	1.0	-	-	-	-	0.0	1.0	0.500	13
Alladian	ald	Volta-Congo	0.50	1.0	0.50	1.0	1.0	0.0	0.0	0.5	-	-	-	0.0	0.5	0.5	-	-	-	-	0.0	1.0	0.500	13
Bobo Madaré	bbo	Mande	0.50	1.0	0.50	1.0	1.0	0.0	0.0	0.5	-	-	-	0.0	0.0	1.0	-	-	-	-	0.0	1.0	0.500	13
Dangaléat	daa	Chadic	0.50	1.0	0.50	1.0	1.0	0.0	1.0	0.5	-	-	-	0.0	0.5	0.5	-	-	-	-	0.0	0.0	0.500	13
Fyem	pym	Volta-Congo	1.00	0.5	1.00	0.5	1.0	0.0	1.0	0.5	-	-	-	0.0	0.5	0.5	-	-	-	-	0.0	0.0	0.500	13
Amo	amo	Benue-Congo	0.50	1.0	0.25	1.0	1.0	0.0	0.0	0.5	-	-	-	0.0	0.5	0.5	-	-	-	-	0.0	1.0	0.481	13
Margi	mrt	Chadic	0.75	0.0	1.00	1.0	1.0	0.0	1.0	0.5	-	-	-	0.0	0.5	0.5	-	-	-	-	0.0	0.0	0.481	13
Ejagham	etu	Benue-Congo	0.50	1.0	0.50	0.5	1.0	0.0	0.0	0.0	-	-	-	0.0	0.5	1.0	-	-	-	-	0.0	1.0	0.462	13
Kanakuru	kna	Chadic	0.50	0.5	0.50	0.5	1.0	0.0	1.0	1.0	-	-	-	0.0	0.5	0.5	-	-	-	-	0.0	0.0	0.462	13
Lelemi	lef	Volta-Congo	0.50	1.0	0.50	1.0	1.0	0.0	0.0	0.5	-	-	-	0.0	0.0	0.5	-	-	-	-	0.0	1.0	0.462	13
Temne	tem	Atlantic-Congo	0.25	1.0	0.25	0.5	0.5	0.0	0.0	0.5	-	1.0	0.75	0.0	1.0	0.0	-	-	-	-	0.0	1.0	0.450	15
Toro So	dts	Dogon	0.25	1.0	0.25	0.5	1.0	0.0	0.0	0.5	-	1.0	0.75	0.0	0.0	0.5	-	-	-	-	0.0	1.0	0.450	15
Aghem	agq	Benue-Congo	0.50	1.0	0.25	1.0	0.5	0.0	0.0	0.5	-	-	-	0.0	0.5	0.5	-	-	-	-	0.0	1.0	0.442	13
Ijo	ijc	Ijoid	0.50	1.0	0.25	1.0	1.0	0.0	0.0	0.5	-	-	-	0.0	0.0	0.5	-	-	-	-	0.0	1.0	0.442	13
Supyire	spp	Volta-Congo	0.50	1.0	0.50	1.0	0.5	0.0	0.0	0.5	1.0	-	-	0.0	0.5	1.0	1.0	0.0	0.0	-	0.0	0.0	0.441	17
Diola-Fogny	dyo	Atlantic-Congo	0.50	1.0	0.00	0.5	1.0	0.0	0.0	0.5	1.0	-	-	0.0	1.0	0.0	1.0	1.0	0.0	-	0.0	0.0	0.441	17
Kera	ker	Chadic	0.50	0.5	0.50	1.0	1.0	0.0	1.0	0.5	-	-	-	0.0	0.5	1.0	0.0	0.0	1.0	1.0	0.0	0.0	0.441	17
Bambara	bam	Mande	0.50	1.0	0.50	1.0	1.0	0.0	0.0	0.5	1.0	-	-	0.0	0.0	0.5	-	-	-	-	0.0	0.0	0.429	14
Kisi	kss	Atlantic-Congo	0.25	1.0	0.25	0.0	1.0	0.0	1.0	0.5	-	-	-	0.0	0.5	1.0	-	-	-	-	0.0	0.0	0.423	13
Dagbani	dag	Volta-Congo	0.50	0.5	0.75	1.0	1.0	0.0	0.0	0.5	-	0.0	0.00	0.0	0.5	0.5	-	-	-	-	0.0	1.0	0.417	15

LANGUAGE	ISO	FAMILY	1A	2A	3A	4A	5A	6A	7A	8A	9A	10A	10B	11A	12A	13A	14A	15A	16A	17A	18A	19A	ICP	n.
Noni	nhu	Volta-Congo	0.50	1.0	0.25	0.5	0.5	0.0	0.0	0.5	-	-	-	0.0	1.0	1.0	-	-	-	0.0	0.0	0.0	0.404	13
Kanuri	knc	Saharn	0.50	1.0	0.50	1.0	0.5	0.0	0.0	0.5	1.0	0.0	-	0.0	0.5	0.5	-	-	-	0.0	0.0	0.0	0.400	15
Hausa	hau	Chadic	0.75	0.5	0.75	1.0	0.5	0.0	1.0	0.5	0.0	0.0	-	0.0	0.5	0.5	-	-	-	0.0	0.0	0.0	0.400	15
Yoruba	yor	Benue-Congo	0.25	1.0	0.25	0.5	0.5	0.0	0.0	0.5	0.0	1.0	-	0.0	0.0	1.0	-	-	-	0.0	1.0	0.0	0.400	15
Klao	klu	Volta-Congo	0.00	1.0	0.00	0.5	0.5	0.0	0.0	0.0	-	-	-	0.0	0.0	1.0	-	-	-	1.0	1.0	0.0	0.385	13
Wolof	wlf	Atlantic-Congo	0.75	1.0	0.50	0.5	1.0	0.0	0.0	0.5	-	0.0	-	0.0	1.0	0.0	-	-	-	0.0	0.0	0.0	0.375	14
Fe'fe'	fmp	Volta-Congo	0.25	1.0	0.25	1.0	0.5	0.0	0.0	0.0	-	-	-	0.0	0.5	1.0	-	-	-	0.0	0.0	0.0	0.346	13
Akan	aka	Volta-Congo	0.50	1.0	0.25	0.5	1.0	0.0	0.0	0.0	-	-	-	0.0	0.5	0.5	-	-	-	0.0	0.0	0.0	0.327	13
Bisa	bib	Mande	0.25	0.5	0.50	1.0	1.0	0.0	0.0	0.5	-	-	-	0.0	0.5	0.0	-	-	-	0.0	0.0	0.0	0.327	13
Efik	efi	Benue-Congo	0.00	1.0	0.00	0.5	0.0	0.0	0.0	0.0	-	-	-	0.0	0.5	0.5	-	-	-	0.0	1.0	0.0	0.269	13

Table A – IPC_{WALS(20)} for 56 WA languages listed in the WALS

#	Language	ISO	IPC _{WALS(4)}
1	Angas	anc	0.938
2	Kohumono	bes	0.917
3	Doyayo	dow	0.875
3	Noni	nhu	0.875
4	Aizi	ahp	0.833
5	Gā	gaa	0.813
5	Bagirmi	bmi	0.813
5	Ewondo	ewo	0.813
5	Ndut	ndv	0.813
5	Kotoko	ktk	0.813
6	Gwari	gbr	0.750

#	Language	ISO	IPC _{WALS(13)}
1	Doyayo	dow	0.654
2	Angas	anc	0.615
2	Gwari	gbr	0.615
2	Kpelle	xpe	0.615
2	Tera	ttr	0.615
3	Kohumono	bes	0.604
3	Aizi	ahp	0.604
4	Kotoko	ktk	0.596
4	Tarok	yer	0.596
5	Bété	bev	0.577
5	Dan	dnj	0.577

#	Language	ISO	IPC _{WALS(20)}
1	Doyayo	dow	0.700
2	Angas	anc	0.615
2	Gwari	gbr	0.615
2	Kpelle	xpe	0.615
2	Tera	ttr	0.615
3	Kohumono	bes	0.604
3	Aizi	ahp	0.604
4	Tarok	yer	0.596
5	Bété	bev	0.577
5	Dan	dnj	0.577
5	Mumuye	mzm	0.577

#	Language	ISO	IPC _{WALS(4)}
6	Kpelle	xpe	0.750
6	Tera	ttr	0.750
6	Bété	bev	0.750
6	Dan	dnj	0.750
6	Mambila	mcu	0.750
6	Konyagi	cou	0.750
6	Biom	bom	0.750
6	Kpan	kpk	0.750
6	Tampulma	tpm	0.750
6	Bariba	bba	0.750
6	Ejaghani	etu	0.750
6	Supyire	spp	0.750
7	Tarok	yer	0.688
7	Ewe	ewe	0.688
7	Isoko	iso	0.688
7	Igbo	ibo	0.688
7	Kisi	kss	0.688
7	Wolof	wlf	0.688
7	Fe'fe'	fmp	0.688
8	Mumuye	mzm	0.625
8	Ogbia	ogb	0.625
8	Ngizim	ngi	0.625
8	Grebo	grj	0.625
8	Alladian	ald	0.625

#	Language	ISO	IPC _{WALS(13)}
5	Mumuye	mzm	0.577
5	Ngizim	ngi	0.577
6	Mambila	mcu	0.563
7	Gā	gaa	0.558
7	Ogbia	ogb	0.558
8	Konyagi	cou	0.556
9	Biom	bom	0.538
10	Bagirmi	bmi	0.519
10	Ewondo	ewo	0.519
10	Ndut	ndv	0.519
10	Kpan	kpk	0.519
10	Tampulma	tpm	0.519
10	Ewe	ewe	0.519
10	Isoko	iso	0.519
10	Igbo	ibo	0.519
11	Bariba	bba	0.500
11	Alladian	ald	0.500
11	Bobo Mad.	bbo	0.500
11	Dangaléat	daa	0.500
11	Fyem	pym	0.500
11	Kera	ker	0.500
12	Amo	amo	0.481
12	Dagbani	dag	0.481
12	Margi	mrt	0.481

#	Language	ISO	IPC _{WALS(20)}
6	Mambila	mcu	0.563
7	Gā	gaa	0.558
7	Ogbia	ogb	0.558
8	Konyagi	cou	0.556
9	Ewe	ewe	0.554
10	Bagirmi	bmi	0.542
11	Biom	bom	0.538
12	Ngizim	ngi	0.536
13	Ewondo	ewo	0.519
13	Ndut	ndv	0.519
13	Kpan	kpk	0.519
13	Tampulma	tpm	0.519
13	Isoko	iso	0.519
14	Kotoko	ktk	0.517
14	Igbo	ibo	0.517
15	Grebo	grj	0.513
16	Bariba	bba	0.500
16	Alladian	ald	0.500
16	Bobo Mad.	bbo	0.500
16	Dangaléat	daa	0.500
16	Fyem	pym	0.500
17	Amo	amo	0.481
17	Margi	mrt	0.481
18	Ejaghani	etu	0.462

#	Language	ISO	IPC _{WALS(4)}
8	Bobo Madaré	bbo	0.625
8	Dangaléat	daa	0.625
8	Fyem	pym	0.625
8	Amo	amo	0.625
8	Aghem	agq	0.625
8	Diola-Fogny	dyo	0.625
8	Kera	ker	0.625
8	Kanuri	knc	0.625
8	Akan	aka	0.625
9	Temne	tem	0.563
9	Hausa	hau	0.563
9	Yoruba	yor	0.563
10	Kanakuru	kna	0.500
10	Lelemi	lef	0.500
10	Ijo	ijc	0.500
10	Bambara	bam	0.500
10	Dagbani	dag	0.500
10	Klao	klu	0.500
10	Efik	efi	0.500
11	Margi	mrt	0.438
11	Toro So	dto	0.438
12	Bisa	bib	0.313

#	Language	ISO	IPC _{WALS(13)}
13	Ejagham	etu	0.462
13	Hausa	hau	0.462
13	Kanakuru	kna	0.462
13	Lelemi	lef	0.462
14	Grebo	grj	0.442
14	Aghem	agq	0.442
14	Ijo	ijc	0.442
15	Supyire	spp	0.423
15	Kisi	kss	0.423
16	Noni	nhu	0.404
16	Wolof	wlf	0.404
17	Kanuri	knc	0.385
17	Temne	tem	0.385
17	Yoruba	yor	0.385
17	Bambara	bam	0.385
17	Klao	klu	0.385
17	Toro So	dto	0.385
18	Fe'fe'	fmp	0.346
18	Diola-Fogny	dyo	0.346
19	Akan	aka	0.327
19	Bisa	bib	0.327
20	Efik	efi	0.269

#	Language	ISO	IPC _{WALS(20)}
18	Kanakuru	kna	0.462
18	Lelemi	lef	0.462
19	Temne	tem	0.450
19	Toro So	dto	0.450
20	Aghem	agq	0.442
20	Ijo	ijc	0.442
21	Supyire	spp	0.441
21	Diola-Fogny	dyo	0.441
21	Kera	ker	0.441
22	Bambara	bam	0.429
23	Kisi	kss	0.423
24	Dagbani	dag	0.417
25	Noni	nhu	0.404
26	Kanuri	knc	0.400
26	Hausa	hau	0.400
26	Yoruba	yor	0.400
27	Klao	klu	0.385
28	Wolof	wlf	0.375
29	Fe'fe'	fmp	0.346
30	Akan	aka	0.327
30	Bisa	bib	0.327
31	Efik	efi	0.269

Table B – IPC_{WALS(20)} vs IPC_{WALS(13)} vs IPC_{WALS(4)}

Sergio Baldi

ORCID 0000-0003-4099-1497

Università degli Studi di Napoli "L'Orientale"

Rudolf Leger

J.W. Goethe-Universität, Frankfurt am Main

Innovative features of nouns and pronouns in Chadic languages of the Nigerian Gongola-Benue basin

Abstract

Many Chadic languages spoken in the Gongola Benue basin of North-Eastern Nigeria have undergone tremendous changes in the domain of their phonology, morphology and lexicon. This in particular concerns the languages Kwami, Kupto, Piya, Widala and Nyam, which are spoken in typical retreat areas by not more than fifteen thousand speakers. The vicinity of neighbouring Adamawa and Jarawan Bantu languages and the influence of Hausa may have led to linguistic innovations in their grammar as well as in the lexicon. The impact can often be proved by the inflation of sound systems, tone levelling in favour of the low tone and the loss of nominal plural, which generally is compensated by verbal pluractionals. The changes may be further observed in an enlargement of pronominal sets, the existence of logophoric pronouns and the use of Intransitive Copy Pronouns in the several languages.

Keywords: linguistic innovations, gender assignment, pluractionals, Intransitive Copy Pronoun, Bole-Tangale languages

1. Introduction

One of the highest ethnic and linguistic density in Africa is represented by the middle belt of Nigeria and particularly the Gongola-Benue basin. Here you find the three greatest African linguistic stocks, i.e. Niger-Congo, Nilo-Saharan and Hamito-Semitic (or AfroAsiatic, Jungrathmayr & Leger 1993: 161).

The close vicinity of the neighbouring Benue-Congo – and here generally of Adamawa- and Jarawan-Bantu languages – has led to considerable changes, so that the Chadic languages display several “innovative and evolutionary features in phonological, morphological and lexical respects”, whose vitality could well be the reflection of a turbulent linguistic history (Jungrathmayr & Leger 1993: 164). The most prominent typological features observed will be described in the following article, where our focus is directed to nouns and pronouns. It should, however, be stressed that we do not only find a certain typological dichotomy in the morphological characteristics of nouns and pronouns, but also in the phonological as well as tonological systems. Therefore a few general remarks on peculiarities of phonology and tonology should be mentioned first.

2. Dichotomy in systemic relations within Chadic at the level of phonology, tonology and morphology

There is an interesting distribution of consonants and vowels among the various languages of the Bole-Tangale group, which somehow goes along with their geographical location as there is a steady increase of consonants and vowels from the northern towards the southern located languages. This also holds for the tones. Nearly all languages exposed to the neighbourhood of Benue-Congo languages in the South show a kind of a ‘tone levelling’ or ‘tone equalizing’ towards a ‘[frozen] low tone’ in most of the cases – be it in the lexicon or in the overwhelming part of the grammar – just contrary to the languages spoken in ‘Chadic’ environment, i.e. in the North and the East, where the flexibility of tone plays an important role in the grammar. Historically interpreted this ‘frozen tone’ in the South seems to be the final stage of a development – especially what concerns the TAM-system of a language – in which all other grammatical possibilities or strategies have already been fully utilized (Leger 2014: 230-234).

If we give a brief look at the morphological typology of the Bole-Tang(a)le languages, then we realise that there is a strong dichotomy as to many parts of the grammar. Languages spoken in the northern part of the Benue-Gongola basin (like Bole, Kwami, Kupto and Maha) are by far more ‘conservative’ in its nominal and verbal morphology, than the southern representatives (like Tangale, Pero, Piya, Widala, Kushi or Nyam). Many ‘southern languages’ have reduced or even lost certain ‘traditional Chadic’ characteristics, which they had before, but at the same time developed a strategy of compensation, which – at least partly – can be considered as being evoked by the surrounding Benue-Congo languages to keep the balance like in ‘communicating vessels’. In the following article we will give a more detailed description of the most striking morphological features displayed in the domain of nouns and pronouns.

3. The interrelationship between gender distinction and plural formation

Let us now have a closer look at the noun in the various languages. All languages of the Bole-Tangale group display grammatical gender without an overt gender marker, which neither morphologically nor tonologically distinguishes masculine from feminine nouns (Leger 2014: 236-237). An interesting fact, however, is that a gender levelling (*Genusnivellierung*) in favour of the feminine gender can be observed especially among the geographically southern languages (Baldi & Leger 2011: 23). This can even lead to a complete loss of gender distinction with only marginal exceptions. The following examples taken from Kwami, Kushi, Piya and Widala may show this phenomenon:

(1) Gloss:	Kwami:	Kushi:	Piya:	Widala:
‘knee’	<i>pùrùm</i> (M)	<i>vùrùm</i> (F)	<i>bùrmi</i> (F)	<i>vùlùm</i> (F)
‘saliva’	<i>ʔillé</i> (M)	<i>ʔèlèt</i> (F)	<i>ʔèléc</i> (F)	<i>ʔèlèk</i> (F)
‘tree’	<i>shòobì</i> (F)	<i>bìró</i> (F)	<i>bùrò</i> (F)	<i>v(ù)rò</i> (F)

The closest related northern language to Kwami is Kupto. In Kupto, where the ratio of feminine to masculine nouns is roughly 90 to 10, no specific semantic fields are found which differentiate masculine from feminine nouns. In Kwami it can be observed that fruits growing in the field are masculine, but when harvested and brought back to the domain of household, they change their masculine gender to feminine (Leger 1998: 206 f.).

In Kushi, a language spoken in close vicinity to Adamawa languages like Bangwinji, Lo and Burak, all nouns with the exception of *kèwù* (M) ‘white beniseed’ and *làdùk* (M), also a type of ‘white beniseed’ are feminine. The explanation given was that these types of beniseed can grow by dew (and are therefore planted in the dry season), whereas all other plants need rain for their growth. Here, a mythological gender assignment must be taken into account.

No explanation was given for the changing gender assignment of *tèrè* ‘moon’. When the moon increases or decreases it is considered as grammatically masculine, but when the moon is full it possesses feminine gender.

And finally, in Piya only one noun with a grammatical gender was found in the whole lexicon material. This is *còmbò* ‘name, name giving ceremony’, which like in Kwami, Kupto and Hausa – in spite of its ending in –aa – carries the masculine gender. See examples:

(2) Kwami	Kupto	Hausa
<i>ʔimi</i> (M)	<i>wín</i> (M)	<i>súnáa</i> (M)

Here one fundamental question arises: why does gender levelling goes in favour of the feminine gender? Presumably – so at least we assume – that all Bole-Tangale languages historically displayed a masculine and feminine grammatical gender marked by specific suffixes. After the loss of such gender marking suffixes, the system undifferentiated the grammatical gender and therefore became less overt. The languages had to decide which basic gender they will follow. Although generally in Afroasiatic the marked category is the feminine gender (i.e. take the Arabic case of *muallim* vs *muallimatun*) it is surprising that languages like Kupto, Maha, Tangale, Pero, Kushi, Piya and Nyam display the feminine gender in spite of the “historical higher marked barrier”. One can also presume (cf. Corbett 1991: 318) that a transfer of principals in the world view has triggered such semantic transfer or that a feminine perception of the world stands as ideology for a basic conception of life (Leger 1998: 208-209; Leger 2014: 238).

If we now compare the findings from a typological point of view, the increase of feminine gender is linked with some other grammatical parameters of the various languages in the nominal and verbal sphere. On the basis of these findings, the following can be stated:

- a) The increase of feminine gender goes along with the decrease of nominal plural formation.

In Kwami, where the distribution of masculine and feminine gender is quite balanced (60 to 40 %), four types of possible plural formation can be found. In Kupto, where the distribution of feminine and masculine nouns is 90 to 10 %, three types of plural formation exists. And in Kushi, Piya, Widala and Nyam, where almost all nouns possess the feminine gender, only one type – and here only the suppletive plural for human beings – is recorded. Other possibilities to form a nominal plural do not exist.

- (3) Kwami
 (SG) *dìndì* – (PL) *dìndì-shíná* ‘story’
 (SG) *yáábé* – (PL) *yéppín* ‘chicken’
 (SG) *kúmíyó* – (PL) *kùmáyà-shíná* ‘co-wife’
 (SG) *zùm* – (PL) *zùm-zùm* ‘difference’
 (SG) *mánú* – (PL) *ômbàrà* ‘woman’
- (4) Kupto
 (SG) *fóhó* – (PL) *fóhó-láy* ‘mouth’
 (SG) *músó* – (PL) *mússén* ‘co-wife; envy’
 (SG) *bíu(-)* – (PL) *‘àná(-)* ‘possessor of’
- (5) Kushi
 (SG) *niṅò* – (PL) *mémé* ‘human being, person’
 (SG) *là* – (PL) *shíbà* ‘child’
 (SG) *péemun* – (PL) *shérèp* ‘wife’
- (6) Piya
 (SG) *míyá* – (PL) *lèenyá* ‘person’
- (7) Widala
 (SG) *míyá* – (PL) *mèmmè* ‘person’
- (8) Nyam
 (SG) *múdùk* – (PL) *sùlúp* ‘woman, female’

b) The increase of feminine gender is connected with the loss of noun plurality

These distributional characteristics correspond proportionally with the increase of verbal plurals (pluractionals after Newman 2012:185-209). In Kwami and Kupto verbal plurality, which is an agreement plural, is expressed by suffix morphemes in the perfective tense only, whereas Kushi and Piya, for example, have developed their own verbal plural stems (VPL), used in all tenses and aspects:

- (9) Kwami
 (VPL): *yìn rèeb-án-gò* ‘they have cleaned’ (SG) *rèeb-ù-gó*
 3PL clean-PLUR-PFV clean-SG-PFV
- (10) Kupto
 (VPL): *mà sháat-à* ‘you have helped’ (SG) *sháat-ù*
 2PL help-PLUR. PFV help-SG. PFV

Examples of the distinction between verbal plural and singular stems in Kushi are given in the respective verbo-nominal form, i.e. (VNPL) *dìjyàni* (SG. *díyàni*) ‘to sit, to settle, to live’; (VNPL) *pàrdàni* (SG. *pàttàni*) ‘to deliver’; (VNPL) *kùṅràni* (SG. *kùràni*) ‘to hate’; (VNPL) *shèyàyàni* (SG. *shèyàni*) ‘to drink’; (VNPL) *shùyàyàni* (SG. *shàani*) ‘to eat’; (VNPL)

yìbìyàni (SG. *yìwàni*) ‘to hold’; (VNPL) *làrràni* (SG. *làttàni*) ‘to hide’; (VNPL) *nìcìyàn* (SG. *nìyàn*) ‘to ripen’;

Similarly, while presenting examples in Piya, we also refer to the verbo-nominal form, i.e. (VNPL) *fùddìrànni* (SG. *fùddànni*) ‘to wash’; (VNPL) *cwàcìrànni* (SG. *cwàrànni*) ‘to decrease’; (VNPL) *wùccìyànni* (SG. *wùyànni*) ‘to dry up’; (VNPL) *làdìlànni* (SG. *làalànni*) ‘to burn’; (VNPL) *cùppùbànni* (SG. *cùbbànni*) ‘to show’; (VNPL) *tàttìrànni* (SG. *tàrànni*) ‘to hide’; (VNPL) *bàngìlànni* (SG. *bàlànni*) ‘to count’.

c) feminine gender and tone levelling

In the languages with an increase of feminine gender additionally a certain tone levelling towards the low tone (*Tonnivellierung*) can be observed. See the following examples from Piya, Kwami, Kupto, and Kushi:

(9)	Gloss	Piya	Kwami	Kupto	Kushi
	‘pregnancy’	<i>pùlè</i>	<i>pùlè</i>	<i>fùlí</i>	<i>ùu</i>
	‘friend’	<i>cèerè</i>	<i>shérè</i>	<i>céré</i>	<i>jèerè</i>
	‘fear’	<i>còwò</i>	<i>shówò</i>	<i>shòwó</i>	<i>shèwò</i>
	‘blind’	<i>pùgùm</i>	<i>pùgùm</i>	<i>fùhùm</i>	<i>fùgùm</i>
	‘finger’	<i>kwáddàkh</i>	<i>kwáddáǵí</i>	<i>kwáddàk</i>	<i>pò khirà</i>
	‘new’	<i>fòyòy</i>	<i>póyó</i>	<i>fóoyò</i>	<i>fò</i>

4. Expanding the pronominal systems

If we now turn to the pronouns and here in particular to the subject pronouns, then quite a contrary development to that of the nouns can be perceived. Here a steady increase or inflation of the various pronominal systems can be observed, absorbing or compensating for the decline or the loss of number and gender markers in the nominal sphere. In addition, it should be emphasised that some of the pronouns are taken as TAM markers.

Four aspects of this phenomenon will be discussed below (cf. a – d):

a) new sets of subject pronouns

In those languages where number and gender are well reflected like in Kwami – and to a certain extent also in Bolanci, Ngamo and Karekare (Dera) – only one pronominal set of subject pronouns is found for all verbal paradigms. In languages where gender levelling towards feminine gender gains the upper hand, like in Kupto, one can find – at least typologically – an extension of subject pronouns in terms of new sets, which are reserved for present continuous, perfect + subjunctive and future. In Piya, Widala and Nyam, as well as in Maaka, where nearly all nouns carry the feminine gender, i.e. where no gender and number distinction in the nominal sphere exist, two or even three pronominal sets can be found. In Piya and Widala we find one additional set for the subjunctive with an extension of an element *-n* and in Nyam one for future, one for present continuous and perfect and one which is exclusively reserved for the subjunctive paradigm. Moreover, Piya and Widala

possess an inclusive and exclusive pronoun in the first person plural (not marked in the table below), a feature which also appears in Kushi.

Examples:

Table 1. Pronominal sets of subject pronouns

	Bole	Kwami	Kupto			Piya			Kare-kare	Ngamo
SG 1.	<i>ɗn</i>	<i>nè</i>	<i>né</i>	<i>nè</i>	<i>nà</i>	<i>né</i>	<i>nè</i>	<i>nèn</i>	<i>nà</i>	<i>né</i>
2. m	<i>ká</i>	<i>kè</i>	<i>ké</i>	<i>kè</i>	<i>ká</i>	<i>kò</i>	<i>kè</i>	<i>kèn</i>	<i>ká</i>	<i>kò</i>
F	<i>shí</i>	<i>shì</i>	<i>shí</i>	<i>shì</i>	<i>cí</i>	<i>shì</i>	<i>cì</i>	<i>cìn</i>	<i>cí</i>	<i>shì</i>
3. m	<i>(íshì) (à)</i>	<i>yì</i>	<i>yí</i>	<i>yì</i>	–	–	<i>càkéy</i>	<i>càkén</i>	–	–
F	<i>(ítá) (à)</i>	<i>tè</i>	<i>té</i>	<i>tè</i>	–	–	<i>tè</i>	<i>tèn</i>	–	–
Pl. 1.	<i>mú</i>	<i>mìn</i>	<i>mín(ù)</i>	<i>mìn</i>	<i>mú</i>	<i>mù</i>	<i>mìn(ù)</i>	<i>mìnùn</i>	<i>mú</i>	<i>mù</i>
2.	<i>māa</i>	<i>mà</i>	<i>má</i>	<i>mà</i>	<i>kú</i>	<i>ngù</i>	<i>mà</i>	<i>màn</i>	<i>kú</i>	<i>ngù</i>
3.	<i>(màté) (à)</i>	<i>yìn</i>	<i>wún</i>	<i>wùn</i>	-	-	<i>cùn(ù)</i>	<i>cùnùn</i>	-	-

Table 2. Pronominal sets of subject pronouns

	Maaka		Widala		Nyam			
Sg. 1.	<i>nì</i>	<i>nàa</i>	<i>nè</i>	<i>nèn</i>	– / ù	<i>ndà</i>	<i>Ndàa</i>	<i>nà(-n)</i>
2. m	<i>kè</i>	<i>kàa</i>	<i>khè</i>	<i>khèn</i>	<i>kì</i>	<i>ká</i>	<i>Káa</i>	<i>nàa-gí</i>
F	<i>shì</i>	<i>shàa</i>	<i>shì</i>	<i>shìn</i>	<i>kyì</i>	<i>kyá</i>	<i>Kyáa</i>	<i>nàa-jí</i>
3. m	<i>sì</i>	<i>yàa</i>	<i>càkhù</i>	<i>càkhùn</i>	<i>nyì</i>	<i>nyà</i>	<i>Nyàa</i>	<i>nàa-nyí</i>
F	<i>tà</i>	<i>tàa</i>	<i>tè</i>	<i>tèn</i>	<i>sì</i>	<i>sá</i>	<i>Sáa</i>	<i>nàa-rí</i>
Pl. 1.	<i>mínè</i>	<i>mùnà</i>	<i>mùnù</i>	<i>mìnùn</i>	<i>ɗán</i>	<i>ɗàndà</i>	<i>ɗàndàa</i>	<i>nàa-án</i>
2.	<i>kù</i>	<i>màa</i>	<i>mè</i>	<i>mèn</i>	<i>kà</i>	<i>kàrà</i>	<i>Kàráa</i>	<i>nàa-gá</i>
3.	<i>kútà</i>	<i>kútàa</i>	<i>cùnù</i>	<i>cùnùn</i>	<i>kày</i>	<i>kàyrà</i>	<i>Kàyráa</i>	<i>nàa-gáy</i>

b) Intransitive Copy Pronouns

Apart from the different pronominal sets, there is another striking feature, namely the occurrence of the so-called Intransitive Copy Pronoun which “copies the person number and gender of the subject onto intransitive verbs” (Newman 1974: 23). This grammatical feature again is found exactly in those languages where gender levelling has taken place. This means – apart from Kwami and restricted in Bolanci – it is found in all other languages of the southern Bole-Tangale group like Kupto, Pero, Piya, Kushi, Widala and Nyam, as well as in Maaha, Tangale and Kanakuru (Dera).

Examples:

(10) Kupto

wùn *rìw-wù* ‘they have entered’
 3PL.PFV enter-ICP.3PL

- (11) Pero
mín -tá -lékkédée-mù ‘we will disperse’ (Frajzyngier 1989: 114)
 1PL-FUT-disperse-ICP.1PL
- (12) Piya
nè ’illigée-nó ‘I have stood up’ (Leger 2014: 242)
 1SG stand.PFV-ICP.1SG
- (13) Kushi
shìnòrì nèsháa-jò ‘they have taken rest’ (Batic, p.c.)
 3PL rest.PFV-ICP.3PL
- (14) Widala
tè mùràn-ró ‘she has died’ (Leger 2014: 242)
 3SG.F.PFV die-ICP.3SG.F
- (15) Maaha
yì ’illò-nì ‘he is standing’
 3SG.M stand-ICP.3SG.M
- (16) Tangale
ambú-no ‘I mount(ed)’ (Jungraitmayr 1991: 59)
 mount-ICP.1SG
- (17) Kanakuru
amboi wù lewo-wu ‘the boys have (get) tired’ (Newman 1974: 23)
 boys 3PL.PFV be tired-ICP.3PL

c) logophoric pronouns

There is a further phenomenon termed logophoricity, which is found in the pronominal systems of Tangale, Kushi, Pero, Piya and Widala (Leger & Zoch 2006: 289f.). And again it should be emphasised that this grammatical feature¹ arises exactly where we find number and genderless languages. The logophoric pronouns may serve here as a kind of substitute for the loss of number and gender in terms of reference tracking. Examples from Tangale, Pero and Widala are taken from Leger (2014: 243):

	no co-reference	co-reference with the speaker	co-reference with addressee
(18) Tangale:	<i>mbeendam</i>	<i>yi</i>	<i>peemo</i> (3M)
	<i>mbaastam ta</i>	<i>paaze</i>	(3F)
	<i>mbiindam yini</i>	<i>piimo</i>	(3PL)

¹ The term logophoric is widely used to refer to syntactic contexts of complements of the verbs of saying. On logophoric systems in Chadic see (Frajzyngier 1985).

(19) Pero:	<i>cakka</i>	-----	<i>peemo/u</i> (3M)
	<i>te</i>	-----	(3F)
	<i>cinu</i>	-----	(3PL)
(20) Widala:	<i>càkhi</i>	<i>càkhù</i>	<i>pèemò</i> (3M)
	<i>mòtò</i>	<i>pèeyè</i>	(3F)
	<i>yèelè</i>	<i>piimè</i>	(3PL)
(21) Kushi	<i>shù</i>	-----	<i>pèemò</i> (3M)
	<i>tà</i>	-----	(3F)
	<i>shimòr</i>	-----	(3PL)

Examples:

(22) Tangale: *mbeendam yimgo ka mbeendam nego ka mbeendam nga kina*
 ‘He (1) thought that he (2) said that he (3) had money’

(23) Pero: *pèemè tók-nà cákkà n wàt-tù áádô róccò-ì*
 ‘[He(1) said to him(2)], he(2) kills animals and he (1) will come and eat the liver’

(24) Kushi: *shù tèka-nð yà mð pèemò wàrò* (Batic, p.c.)
 ‘He(1) told me that he(1) should come’

(25) Widala: *càkhù tèkkòw càkhù/càkhi wèeyòw ’indè*
 ‘He(1) said, he(1)/he(2) saw a dog’
 (“anti-logophoric”)²

An interesting aspect which should be mentioned here is that in Tangale the logophoric pronouns which show co-reference with the speaker are the old inherited or classical preverbal subject pronouns in Chadic reconstructed by Newman and Schuh (1974) as **si* and **ta*.

d) two pronominal sets of direct object pronouns

Finally, we want to mention the functions of the direct object pronouns in these languages. At least in Kwami, Kupto and Maaka – and most probably as well in Piya, Widala and Nyam we find two different pronominal sets. The first set is employed in paradigms of non-perfective aspect and looks – at least phenotypically – like the possessive pronouns in the respective languages. The second set is employed exclusively in forms of the perfective aspect.

² Widala is the only one language so far, where an anti-logophoric system has been discovered, i.e. when repeating the pronoun the same person is meant.

Examples:

(26) Kwami

tè shùm-à-n-gò ‘she will beat you (M)’
3SG.F beat-FUT-CON-2SG.M.OBJ

tè shùm-ì-gá ‘she has beaten you (M)’
3SG beat-PFV-2SG.M.OBJ

(27) Kupto

né màad-ò-tò ‘I am telling her’
1SG tell-CONT-3SG.F.OBJ

né màad-ù-tà ‘I have told her’
1SG tell-PFV-3SG.F.OBJ

(28) Maaka

nii mòy-tò ‘I will see her’
1SG.FUT see-3SG.F.OBJ

nì mòy-ti tà ‘I have seen her’
1SG see-PFV 3SG.F.OBJ

5. Summary

Chadic languages – and in particular those bordering Benue-Congo languages in the Nigerian middle belt – have undergone “tremendous transformational processes” in the past. This holds true not only for major, but also for minority languages of the West-Chadic group like Tangale, Kwami, Kupto, Piya, Maaha, Nyam and Widala. All of these languages are spoken in typical retreat areas, by generally not more than 15.000 speakers. The exceptions are Bole and Tangale, whose number of speakers may reach about 50.000 to 70.000 souls. This fact of being isolated but also in the close vicinity of neighbouring Adamawa and Jarawan-Bantu languages has led to certain linguistic innovations and peculiarities in various grammatical domains. Among the most prominent typological features observed are the loss of nominal plural formation and a tone levelling in favour of the low tone (esp. seen in Piya and Widala). But the loss in the realm of the noun leads – like in ‘communicating vessels’ – to an expansion of the pronominal (and verbal) system. Thus, we can observe the appearance of the so called ‘Intransitive Copy Pronouns’ (like in Kupto, Piya, Tangale, Kanakuru, Maha and Widala); the existence of ‘Logophoric Pronouns’ (in Piya and Widala) and the enlargement of the pronominal sets for future and subjunctive (so in Kupto, Piya, Widala and Nyam respectively).

Abbreviations

1,2,3 – 1 st , 2 nd , 3 rd person	PL	plural
CON connector (linker)	PLUR	pluractional (verbal plural)
F feminine	PFV	perfective
FUT future	VNPL	verbo-nominal pluractional (verbo-nominal plural)
ICP intransitive copy pronoun	VPL	verbal plural (stem)
M masculine		
OBJ object		

References

- Corbett, Greville G. 1991. *Gender*. Cambridge/New York/Port Chester/Melbourne/Sydney: Cambridge University Press.
- Crozier, David H. and Roger Blench (ed.). 1992^{2nd}. *An Index of Nigerian Languages*. Abuja, Nigeria and Dallas: Nigerian Language Development Centre, Department of Linguistics and Nigerian Languages, University of Ilorin, and Summer Institute of Linguistics.
- Dinslage, Sabine and Rudolf Leger. 1996. "Language and migration. The impact of the Jukun on Chadic speaking groups in the Benue-Gongola basin." In *Proceedings of International Symposium SFB 268, "Kulturentwicklung und Sprachgeschichte im Naturraum Westafrikanische savanne"*, Frankfurt/ Main. [Berichte des Sonderforschungsbereichs 268, Vol. 8, 67-75.
- Frajzyngier, Zygmunt. 1985. "Logophoric systems in Chadic". *Journal of African Languages and Linguistics* 7, 23-37.
- Frajzyngier, Zygmunt. 1989. *A Grammar of Pero* [Sprache und Oralität in Afrika, Vol. 4]. Berlin: Dietrich Reimer.
- Greenberg, Joseph Harold. 1950. "Studies in African linguistic classification: IV Hamito-Semitic". *South-western Journal of Anthropology* 6, 47-63.
- Jungraithmayr, Herrmann. 1992/93. "On vowel systems of Chadic: a typological overview". *Folia Orientalia* 29, 119-129.
- Jungraithmayr, Herrmann, in collaboration with Njeno Andirya Galadima and Ulrich Kleinewillinghöfer. 1991. *A Dictionary of the Tangale Language (Kaltungo, Northern Nigeria) with a Grammatical Introduction* [Sprache und Oralität in Afrika, Vol. 12]. Berlin: Dietrich Reimer.
- Jungraithmayr, Herrmann & Rudolf Leger. 1993. "The Benue-Gongola-Chad Basin – Zone of ethnic and linguistic compression". In: *Berichte des Sonderforschungsbereichs 268*, Vol. 2, Frankfurt am Main, 161-172.
- Kleinewillinghöfer, Ulrich. 1990. "Aspects of vowel harmony in Waja and Tangale-Waja common vocabulary". *Frankfurter Afrikanistische Blätter* 2, 93-106.
- Kleinewillinghöfer, Ulrich. 1996. "Die nordwestlichen Adamawa-Sprachen: eine Übersicht". *Frankfurter Afrikanistische Blätter* 8, 81-103.
- Leger, Rudolf. 1998. "Grammatical gender in some southern Bole-Tangale languages – Kwami, Kupto, Kushi and Piya". In: Igor Sledzewsky & Dmitri Bondarenko (eds.). *Afrika: Obschestva, kul'tury, yazyki [Africa, Societies, Cultures and Languages]*. Moscow: Institute of African Studies 204-216.
- Leger, Rudolf. 2014. "Some observations on typological features in languages of the Bole-Tangale group". In: Anne Storch, Johannes Harnischfeger & Rudolf Leger (eds.). *Fading Delimitations. Multilingual Settlements in a Convergence Area – Case Studies from Nigeria* [Series: Topics in Interdisciplinary African Studies, Vol. 34]. Köln: Rüdiger Köppe Verlag, 229-261.
- Leger, Rudolf & Anne Storch. 1999. "Zur Genese komplexer Verbalsysteme in einigen nordostnigerianischen Sprachen". *Afrika und Übersee* 82, 179-190.
- Leger, Rudolf & Ulrike Zoch. 2006. "Logoforic pronouns in the southern Bole-Tangale languages". In: Petr Zima & Bernard Caron (eds.). *Sprachbund in the West African Sahel*. Leuven-Paris: Peeters, 205-214.

- Lukas, Johannes. 1934. "Die Gliederung der Sprachenwelt des Tschadsee-Gebietes in Zentralafrika". *Forschungen und Fortschritte* 10(29), 356-57.
- Newman, Paul. 1974. *The Kanakuru Language*. Leeds: Institute of Modern English Language Studies, University of Leeds.
- Newman, Paul. 2012. "Pluractional verbs: An overview". In: Patricia Cabredo Hofherr & Brenda Laca (eds.). *Verbal Plurality and Distributivity*. Berlin: de Gruyter, 185-209.
- Newman, Paul & Roxana Ma. 1966. "Comparative Chadic: Phonology and Lexicon". *Journal of African Languages* 5, 218-51.
- Newman, Paul & Russell G. Schuh. 1974. "The Hausa aspect system". *Afroasiatic Linguistics* 1(1), 1-39.
- Storch, Anne. 1999. *Das Hone und seine Stellung im Zentral-Jukonoid*. Köln: Rüdiger Köppe Verlag.
- Tuller, Laurice Anne. 1997. "Les 'ICP' en tchadique". In: Anne Zribi-Hertz & Charlotte Galves (eds.). *Les pronoms: morphologie, syntaxe, et typologie*. Vincennes: Presses Universitaires de Vincennes, 213-229.

Methodological and technical challenges of a corpus-based study of Naija

Abstract

This paper presents early reflections on the NaijaSynCor survey (NSC) financed by the French *Agence Nationale de la Recherche*. The nature of the language surveyed (Naija, a post-creole spoken in Nigeria as a second language by close to 100 million speakers) has induced a specific choice of theoretical framework (variationist sociolinguistics) and methodology (a corpus-based study using Natural Language Processing). Half-way through the 4 year-study, the initial methodological choices are assessed taking into account the nature of the data that has been collected, and the problems that occurred as early as the initial stages of their annotation.

Keywords: Atlantic pidgins and creoles, corpus studies, natural language processing, syntax, prosody

1. Introduction

The NaijaSynCor survey (NSC¹) is a corpus-based survey of Naija, a pidgincreole (Bakker 2008) spoken in Nigeria as a second language by close to 100 million speakers in Nigeria and in the Nigerian diaspora. The nature and size of the language has compelled us to make the annotation process as automatic as possible, with the help of multiple programmes: PRAAT (Boersma & Weenink 2013) for alignment, Elan-Corpa (Chanard 2014) for transcription-translation and semi-automatic tagging; SPPAS (Bigi & Hirst 2012; Bigi et al. 2017) for phonetisation and syllabification; Analor (Avanzi et al. 2008) for prosodic annotation; Arborator² for dependency syntactic annotation; Trameur (Fleury & Zimina 2014) and Grew (Guillaume et al. 2012) for error mining, information retrieval and analysis. The metadata was processed through an application based on Arbil³, and developed by Christian Chanard (Llacan) to make it more user-friendly.

The parallel use of so many different applications requires precise coordination and constant review to adapt the procedures and ensure a smooth workflow. One of the main

¹ A Corpus-based Macro-Syntactic Study of Naija (Nigerian Pidgin) – NaijaSynCor. *Agence Nationale de la Recherche*. February 2017-July 2020. <https://anr.fr/Project-ANR-16-CE27-0007>. Principal Investigator: Bernard Caron, CNRS-LLACAN.

² <https://arborator.ilpga.fr/>

³ <http://explorationdecorpus.corpusecrits.huma-num.fr/arbil/>

challenges is to make sure that the temporal indexes of the annotations are preserved by the various programmes so that Prosodic, Communicative and Syntactic hierarchies can be projected on each other. Another challenge is linked to the nature of the linguistic object itself: as a rapidly expanding pidgincreole, it is somehow unstable, and we have to deal with innovations for which the annotation system must be revised and revisable without having to redo the annotation and without loss of information. A certain degree of lability must be built into the methodology to allow for this inherent dimension of the research project. Two examples of change of procedure will be given concerning phonetization and syllabification on the one hand, and dependency syntax on the other hand.

This paper concentrates on the methodology, the edition and the annotation of the corpus.

2. Naija and Nigerian Pidgin

Nigeria, with 160 million inhabitants, is a huge and complex multilingual community with over 500 different languages (Lewis et al. 2013) used within the public and private social spaces. Among those, **Nigerian Pidgin**, is spoken as a first language by 5 million people, while over 70 million people use it as a second language or as an interethnic means of communication in Nigeria and in Nigerian Diaspora communities. Since the independence of Nigeria in 1960, this variety of Nigerian Pidgin has been rapidly expanding from its original niche in the Niger delta area to cover two-thirds of the country, up to Kaduna and Jos, and is now deeply rooted in the vast Lagos conurbation of over 20 million people. Apart from its original location and one Lagos district, where it is learnt as a first language and can be used as a single language (Elugbe & Omamor 1991), this emerging variety is learnt alongside and not instead of other Nigerian languages. It has become, over the last 30 years, the most important, most widely spread, and perhaps the most ethnically neutral *lingua franca* used in the country today.

The origin of Nigerian Pidgin itself (NP) is generally described as a development out of an English-lexified jargon attested in the 18th century in the coastal area of the Niger delta (River State), with lexical and structural influence from Krio through the activities of missionaries from Sierra Leone (Faraclas 1996; Huber 1999). Today, the heartland of NP is the Niger Delta, with Lagos and Calabar as secondary extensions. But a new development has taken place over the last 50 years whereby NP has escaped from its original geographical niche, where it functioned as an auxiliary medium of communication in restricted informal contexts by uneducated people (Deuber 2005), and is now commonly used all over Nigeria by the educated in informal conversations, and in formal domains, *viz* radio, television, politics, advertising, Christian religious activities, etc.

This variety of Nigerian Pidgin is gradually becoming a pidgincreole which we call **Naija**⁴ to distinguish it from Nigerian Pidgin, the creole spoken in e.g. Warri, Sapele and the Ajegunle district of Lagos. This paper and the NaijaSynCor project are dealing specifically with Naija.

⁴ *Naija*, based on the etymon *niger* which gave its name to the river, is the term used by NP speakers to refer to *Nigeria*.

In terms of functional status, English is Nigeria's official language, and it is dominant in the education system and in written usages (literature, press, etc.). However, Naija has made considerable progress in formal contexts such as information transmission by government and non-government agencies, Christian religious practices, and although it is still excluded from the educational system, it is used unofficially in multilingual schools in southern Nigeria. Naija is a lingua franca in public informal communication in the south and to a certain extent in the north, and it is noticeably popular among university students and among educated speakers in private informal communication (Egbokhare 2004). Recently, Naija has become ubiquitous on local FM radio stations, and has become the single medium of the Wazobia radio and TV and of the Pidgin BBC station since its launching in August 2017. Last but not least, its use is an identifying feature of Nollywood, the prosperous Nigerian film industry now known all over the world.

At the same time as it grows in terms of status and functions, Naija expands geographically, and it is exposed to vernacular languages belonging to different genetic and typological groups (such as Yoruba in the southwest; Igbo, etc. in the southeast; Hausa further north). In the process, does it undergo some degree of contact-induced variation beyond the odd word borrowed from those vernacular languages, or on the contrary, does one standard variety emerge through the influence of modern mass-media such as radio, television and video?

In its functional expansion, Naija is subject to extensive contact and influence from its original lexifier, i.e. English, which is the dominant formal and official language in Nigeria. A question arises as to the extent of this influence today, and what can be deduced of the future of Naija. Does Naija, despite the influence of English (and the indigenous languages of Nigeria) maintain its existence as a discrete language (Deuber 2005) or is it undergoing "decreolization", resulting in what has been described as a post-creole (P/C) continuum (Rickford 1987)? In such a process, a whole range of "mesolectal" varieties create a continuum between the "basilect" (*viz.*, in our case, Delta NP) and the acrolectal varieties deeply influenced by the original lexifier and its local variant (*viz.* the Nigerian variety of English). Is Agheyisi (1984) right when she states that "the possibility of a systematic mesolectal variety emerging in the Nigerian situation is rather remote" (p. 230)? Deuber (2005) convincingly argues that the Naija variety spoken by educated speakers in Lagos is a discrete language, distinct and separate from English, and "the more competent a speaker is in both languages, the better he/she is able to keep them apart" (p. 203). However, the question remains whether this situation applies to Naija outside Lagos where it is further influenced by local native languages (e.g. Yoruba, Igbo, Hausa).

The influence of written Nigerian English on Naija needs special consideration. The extension of Naija to formal usages such as the radio news report, political and information podcast blogging, Bible translation, short story writing, exposes the language to the influence of written Nigerian English. News reports on the radio are generally translated from press releases issued in English by news agencies. Podcast blogs are read from written texts. This new dimension is bound to influence the structure of the language. From the structure of oral Naija, where utterances and information units are mainly structured by information structure, the structure of sentences in written Naija tends to be informed by microsyntax.

3. Objectives and hypotheses

The general aim of the NSC project is to take an exhaustive and in-depth look at the **nature and functions of Naija (Nigerian Pidgin) in Nigeria today**, in order to establish the link between change in structure and change in language use and function. It makes use of the most advanced developments in corpus studies and natural language processing, which combines with a sociolinguistic and geographical study of variation according to formal/informal uses, gender and education of speakers. The corpus studies natural (non-elicited) speech in order to evaluate the distance between Naija and Nigerian English through the study of intonation, information structure, morphology, micro- and macro-syntax.

The distinction between micro- and macrosyntax was first proposed by Blanche-Benveniste et al. (1990), Berrendonner (1990), and Cresti (2000) (but see also Andersen & Nølke (2002) for an overview). These studies put forward macrosyntax as a level of linguistic description capable of accounting for a number of cohesion mechanisms particularly frequent in spontaneous spoken language, which cannot be simply regarded as microsyntactic government phenomena, such as, for example, the “paratactic” constructions in (1) where no conjunction expresses the syntactic link between *you carry your children go* and *you go still buy food*:

1. [*you carry your children go*] [*you go still buy food*] (Deuber 2005)
 [*you bring your children*] [*you will still buy food*]
 ‘[Even if] you bring your children, you will still have to buy food.’

Macrosyntactic models characterize some major linguistic units that go beyond government proper and are usually described in the literature from a pragmatic perspective that focuses on their illocutionary or rhetorical values. Macrosyntax, instead, focuses on the span and the form of macrosyntactic units, using syntactic and distributional criteria (such as suppressions, insertions, commutations) to identify and delimit them. For all macrosyntactic models, the main identifying criterion of a macrosyntactic unit is the possibility that this unit has to constitute an autonomous utterance.

The problems facing any programme of an exhaustive and in-depth study of Naija are many. The first one is related to the popular view of Naija as a protean, ever changing, informal medium that has no unity, and varies with every place and situation where it is spoken. The NSC project is based on the assumption that it is a discrete language with a strong unity that accommodates a certain range of variation.

The second one is related to the success of the Bickerton-DeCamp theory of the creolization-decreolization cycle (cf. above) informing the work of researchers such as Faraclas (1996), Elugbe and Omamor (1991) who approach the study of Naija with a purist attitude, for whom the only form of NP worth studying is the Warri-Sapele “creole” variety spoken in Delta State, and who consider other varieties at best as degraded forms working as a lingua franca commonly called Broken, at worst as “pseudo-pidgins” invading the press and media. Their descriptions of Naija are monolectal, based on their intuition as speakers of the language. The NSC project is data driven and multilectal.

The third one is the difficulty of combining a structural approach with a sociological one. The structural approach, best illustrated by Faraclas (1996) concentrates on the

grammar and vocabulary of the language as revealing the inner mechanisms responsible for the birth and evolution of creoles, pidgins, and languages in general. The sociological approach mostly favoured by Nigerian scholars centres on the study of language usage and representation among speakers (e.g., Ajibade et al. 2012), but the link with the nature and structure of the language is often absent. The NSC project bridges the gap between mental representations of Naija and actual linguistic usages. It plans to use the most recent trends of data-driven, corpus-based sociolinguistics, combining qualitative and quantitative corpus methods with functional and structural analyses.

The fourth one is the difficulty attached to the mere size of the language, and the challenge it represents for a study to account for the geographical and functional variations of a language with millions of speakers. This calls for careful corpus planning and well organised team work, but most of all, it requires the use of NLP tools to make corpus annotation as automatic as possible.

To sum up, the objectives of the NSC project are:

1. Building a reference 500,000 words **oral corpus** (the **Reference Naija Corpus, RNC**), collected in 10 different points of survey in the country, with a deeply annotated sub-section of 100,000 words (the **Naija prosodic and syntactic Treebank, NTB**). Annotated corpora are rare for most of the languages of the world, all the more so if one considers depth of annotation (part-of-speech tagging, syntactic parsing, prosodic annotation). This synchronic picture of Naija, documenting its geographic and demographic variation, is a rare opportunity to study the evolution of a fast emerging, vast new language spoken by tens of millions. This corpus is expected to provide the basis for the standardisation and development of the language.
2. Comparing the RNC with the Nigerian International Corpus of English, ICE Nigeria (Gut 2014), both qualitatively and quantitatively. Naija has been proved to be, in the use of the educated Nigerians living in Lagos, a discrete language that is developing and keeping its own distinctive identity and status separate from English (Deuber, 2005). This study aims to assess whether this holds true in the other parts of Nigeria where it is spoken. This comparison aims at evaluating the discreteness and independence of Naija in relation to Nigerian English, and tests the correlation of potential variations to sociological/functional factors.
3. Achieving a better understanding of the variations of Naija along the formal-informal functional scale through the study of its use on university campuses and in the media, and more specifically on the radio (news reporting, editorials, information, etc.). The following hypotheses will be tested: (i) educated Naija is more standardized possibly due to the geographical and social mobility of its speakers; (ii) it reveals a greater influence from English with more borrowing and more syntactic restructuring; (iii) scripted oral Naija reveals an even stronger influence from English than unscripted oral Naija. The results of the project are expected to provide the basis for the standardisation and development of the language. This assessment of the role and impact of new media in relation with the change of attitude of speakers concerning an emerging language is an unprecedented endeavour. The framework chosen is that of the variationist sociolinguistics framework (Tagliamonte 2012).
4. Understanding the patterns observed in the prosody of emerging languages, and linking the prosodic description of Naija to that of its grammatical and information

structures through the use of NLP tools. The aim is threefold: (i) produce a prosodic description of an underdescribed language in Africa based on instrumental analyses and validated with a resynthesis tools; (ii) provide the Naija Treebank with an in-depth integrated annotation for part-of-speech (POS), intonation, micro- and macro-syntactic structures and information structure, thus producing a gold-standard benchmarking large treebank database, a first for an emerging language; (iii) developing Natural Language Processing (NLP) tools for Naija, namely a POS tagger, an English glosser, and a syntactic parser, integrating a treatment of macrosyntactic constructions (dislocation, clefting ...) and phenomena specific to spoken languages (disfluency, reformulation, discourse markers). The integration of macrosyntax into a syntactic parser is a ground-breaking endeavour, where high-gain results are expected for the development of NLP tools.

4. Methodology

This section retraces the workflow of the NSC project from data collection to editing, annotation and sociolinguistic analysis (Figure 1).

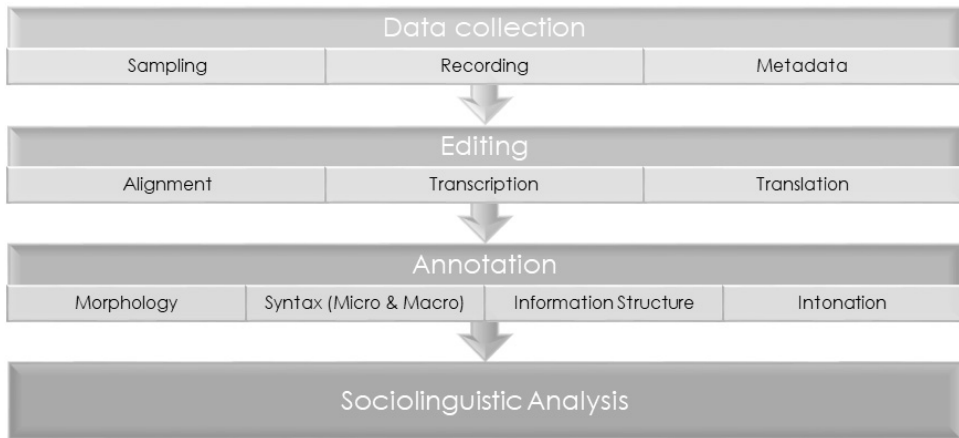


Figure 1. NSC Workflow

4.1. Sampling and data collection

In the absence of sufficient written data, it has been decided to test our hypotheses on an oral corpus to be compared with existing corpora of English, both British and Nigerian, and the Deuber corpora (Deuber 2005) recorded in Lagos almost 20 years ago. The size and nature of the object we want to study imply numerous constraints in collecting the oral data, its annotation, and its analysis.

Due to logistic constraints, our survey focused on urban areas, and more specifically on 10 locations: Lagos, Ibadan, Benin City, Enugu, Onitsha, Port Harcourt, Abuja, Kaduna, Jos, Kano⁵.

⁵ Calabar, which was originally included in the sampling, was left out for logistic reasons, while Onitsha and Enugu could be included in the survey.

Anyone who volunteered and felt confident in their competence in Naija was recorded speaking in monologues and dialogues on any topic they chose themselves. Out of more than 330 speakers recorded, only 2 requested to remain anonymous. Some life stories were collected, as well as 2 speeches on road safety by government agents and 2 sermons from pastors. Radio call-in programmes, news readings and commentaries were recorded in radio stations in Lagos, Ibadan, Kano, Kaduna, Abuja and Port Harcourt. Two excerpts from a drama representing the Passion of Christ staged in Ibadan for Easter 2017 were read by members of the troupe.

The variationist analysis we want to perform implies collecting samples representing different types of speakers, and different types of functions, with a metadata questionnaire documenting the time and place of the interview and the linguistic biography of the speakers (Figure 2). The questionnaires were processed through Arbil with an interface developed by Christian Chanard (Figure 3). After normalisation of data and georeferencing, the metadata can be analysed and mapped to visualise the sampling. (Figure 4).

Dial: 1187yaka: 9JA. ABI. GWA - #7. #9

NAUA – Sociolinguistic Questionnaire FILE N° 7

Wetin be your name? MAZIND DICKSON - AMAGADA

1. You gree for dis recording?
I agree to participate in the recording conducted by the NAUISynCoM Project
at GWAGWALADA
We fit use your name for di recording? Yes: No:
You gree say make researchers dem for use your recording? Yes: No:
Signature: Amagada

2. Recording metadata
Date of recording: JUNE 5TH, 2017
Place of recording: GWAGWALADA
Type of recording
 monologue dialogue radio broadcasting
 other: _____
Other participants: file(s) n° 6 David.
Remarks: _____

3. Personal information: Wetin be your age?
Age group: Under 15 16-30 31-45 46-60 over 60
Sex: male female
Place of birth Where de for born you?
Town/Village: BENIN-CITY State: EDO STATE
Place of residence (if different from the place of birth) Where you stay?
Town/Village: ABUJA State: FCT
When you start to dey stay here?: 2016
How you go school reach?
 Informal primary secondary undergraduate graduate other: _____
Wetin be your work? I BE JOURNALIST (TALK-TALK DEM-SAY)
Remarks: NA CORRECT WAKA BE DIS (Vow)

4. Linguistic profile

Which language you first sabi speak for small pikin?
L1: ENGLISH L2: PIDGIN L3: _____ L4: _____

Which language your papa dey speak?
L1: 1500 L2: ENGLISH L3: PIDGIN L4: _____

Which language your mama dey speak?
L1: YORUBA L2: ENGLISH L3: PIDGIN L4: 1500

Which language (dem) dey take teach une for school?
L1: ENGLISH L2: _____ L3: _____ L4: _____

E get any oder language wey you learn outside house and school?
L1: EDO (BIA) L2: _____ L3: _____ L4: _____

Make you arrange your language as you sabi dem resch?
L1: ENGLISH L2: PIDGIN L3: _____ L4: _____

How you sabi speak Pidgin reach?
 small a dey manage well well well

Remarks: _____

5. Language Practice

Which language you dey use talk to your papa and mama?
L1: ENGLISH L2: PIDGIN L3: _____ L4: _____

Which language you dey use talk to your broder and sister?
L1: ENGLISH L2: PIDGIN L3: _____ L4: _____

Which language you dey use talk to your pikin dem?
L1: ENGLISH L2: _____ L3: _____ L4: _____

Which language you dey use talk to your area people?
L1: PIDGIN L2: _____ L3: _____ L4: _____

Which language you dey use talk to di people wey you follow work?
L1: ENGLISH L2: _____ L3: _____ L4: _____

Which language you dey use talk for any government office, bank, police station?
L1: ENGLISH L2: _____ L3: _____ L4: _____

Remarks: _____

Figure 2. Sample of a metadata questionnaire

NaijaSynCor Session & Speakers Metadata

add SPK > Choose a Session **ABJ_GWA_02_D** > Export sessions to CSV
 Choose a Speaker > > Export speakers to CSV

Search on the database

audioFile	ABJ_GWA_02_D
ELAN File	ABJ_GWA_02_D_Market-Food-Church_T
Date	2017-06-08
Title	At the market
Location	Gwagwalada (FCT)
Genre	Dabigwe
Words	1166
Speaker 1	ABJ_GWA_SP_02_Charity Ohore < DelSPK
Speaker 2	ABJ_GWA_SP_01_Esther Omorugbe < DelSPK
<input type="button" value="Update"/>	

[ABJ_GWA_03_M] [ABJ_GWA_02_D]	[ABJ_GWA_02_D]
-----------------------------------	------------------

<p>Name Charity Ohore Code ABJ_GWA_SP_02 Anon false Words 0 MetaFile ABJ_GWA_SP_02.pdf Sex F Age 16-30 BirthPlace Saloun Gari (Niger) Residence Gwagwalada (FCT) Since 16 yrs Education Secondary Occupation Hair dresser Birth Language Pidgin, Isoko, Hausa, English Father Language Isoko, Hausa, Pidgin, English Mother Language Pidgin, Isoko, Hausa, English School Language English, Hausa, Pidgin Outside Language English, Pidgin, Hausa Language hierarchy Pidgin, Hausa, Isoko, English Practice w/Parents Pidgin, Isoko Practice w/Siblings Pidgin Practice w/Children Pidgin Practice w/Area Pidgin Practice w/Work Pidgin Practice w/Official English, Pidgin Description Remarks ABJ_GWA_02: SP1</p>	<p>Name Esther Omorugbe Code ABJ_GWA_SP_01 Anon false Words 0 MetaFile ABJ_GWA_SP_01.pdf Sex F Age 31-35 BirthPlace Lokoja (Kogi) Residence Gwagwalada (FCT) Since 3 yrs Education Primary Occupation Selling vegetables Birth Language Edo, Pidgin Father Language Edo, Pidgin Mother Language Edo, Pidgin School Language Pidgin Outside Language Yoruba, Hausa Language hierarchy Pidgin, Edo, Hausa Practice w/Parents Edo, Pidgin Practice w/Siblings Pidgin Practice w/Children Pidgin Practice w/Area Pidgin Practice w/Work Pidgin Practice w/Official Pidgin Description Remarks ABJ_GWA_02: SP2 35 yrs old</p>
--	---

Figure 3. A metadata questionnaire processed through Arbil

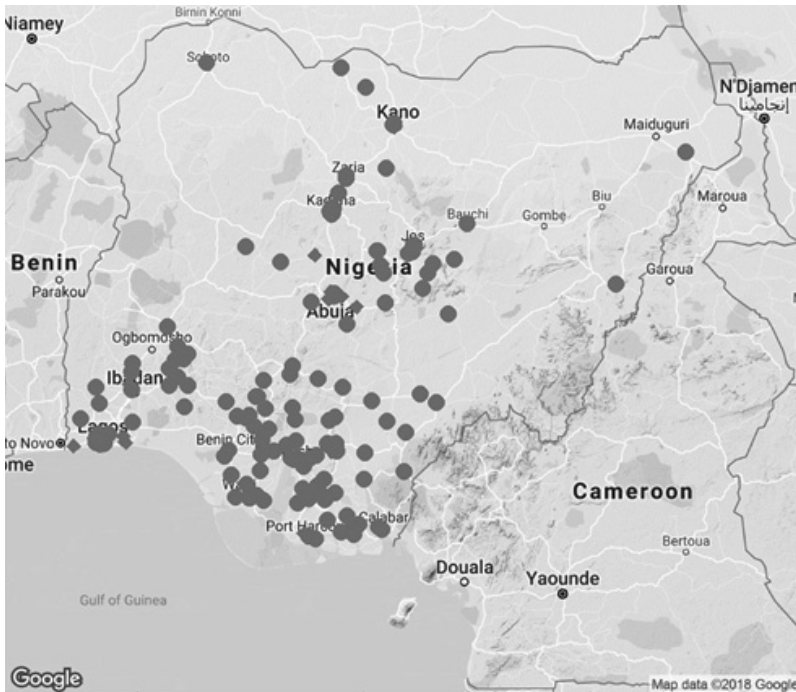


Figure 4. Speakers' birthplaces (Čeplö & Manfredi 2019)

4.2. Annotation, querying and NLP tools

NLP tools are used for corpus editing, morphosyntactic tagging (Elan; §4.3.1); phoneticising (SPPAS) and prosodic modelling (Analog; §4.3.2); syntactic annotation (MATE and Arborator; §4.3.3); error mining, querying and sociolinguistic analysis (Grew and Trameur; §4.4.4).



Figure 5. NLP tools in NSC

4.2.1. Editing the data

Elan (<https://tla.mpi.nl/tools/tla-tools/elan/>) (Sloetjes 2014) was used in annotate the files by providing time alignment, transcription, tokenization into words, semantic word-level glosses and translation into Standard English. Compatibility of transcription has been ensured by using the orthography developed in (Deuber 2005). This etymological orthography (adapted from the lexifier language orthography, i.e. English) has been chosen by Deuber preferably to the phonological script used by linguists (e.g. Faraclas, Elugbe, etc.) as it is spontaneously used by educated Nigerians, and thus easier to teach to transcribers. Codeswitched sections were identified by dedicated boundaries. Transcriptions and translations are double checked for the sake of consistency. A macrosyntactic punctuation marks macro-syntactic boundaries (i.e. illocutionary units and their main components: nucleus, prenuclei and post nuclei, including discourse markers) and limits between pile layers (disfluencies, reformulation, coordination). All these boundaries are marked by punctuations in written texts.

Our segmentation is based on a long tradition of the study of syntax of spoken production in Romance languages (Blanche-Benveniste et al. 1990; Cresti 2000; Simon & Degand 2011). Our maximal syntactic units are illocutionary units, that is, assertions, questions, and demands. We use the markup developed in the Rhapsodie project of annotation of spoken French (Deulofeu et al. 2010; Kahane & Pietrandrea 2012), which is a kind of formalized punctuation. The delimiter for illocutionary units is //. Consider this extract (2) from a sample illustrating the markup:

2. *den you go dey wrap dat food { small |r small } // cut cocoyam // = cut dat uh & // take { cocoyam |c and yam } wey you don grind // = [...]*
 ‘then you will wrap that food in small pieces, cut the cocoyam, cut that er... take the cocoyam and yam which you have ground [...]’ [DEU_A05]

The notation $\{ X | Y \}$ indicates that the phrase Y occupies the same syntactic position as X and piles up on X (Gerdes & Kahane, 2009). Four types of lists are considered: “|c” marks coordination (*cocoyam |c and yam*); “|r” marks (syntactic) reduplication (*small |r small* ‘very small’); “|a” marks appositions (*John |a my friend*); and “||” marks disfluencies and reformulation (*some || some people dey ask* ‘some... some people are asking’).

Inserting the macrosyntactic annotation into the text is part of the segmentation of the transcription and constitutes a first coarse-grained syntactic analysis. The macrosyntactic annotation can be studied as such to quantify phenomena that are more typical for spoken language such as left and right dislocations and disfluencies. It is also geared for the direct study of the prosody-syntax interface. The macrosyntactic annotation improves parsing results and it can easily be simplified into a standard punctuation.

4.2.2. Prosodic analysis

The main objective of this part of NSC is to include a prosodic level in the description of Naija. It will produce (i) an analysis of its prosodic units and their nature, with a description of their precise acoustic correlates, based on an instrumental analysis and validated with a speech synthesis tool; (ii) a version of the 100Kw Golden Corpus annotated for prosody, adapting schemes developed in the treebank Rhapsodie for French, based on perceptual and acoustic cues, developed independently from the micro/macro-syntactic parsing and labelling, which will serve in the final functional analyses. It will answer questions pertaining to the prosodic system of Naija, including phenomena such as speech rhythm, tonal structure, intonation and stress, e.g. (i) Is Naija a ‘tone language’, ‘pitch-accent language’ or ‘stress language’ (Hyman 2006)? (ii) What is the interplay between putative tone and intonation?

Methodology: from the point of view of phonetic instrumentation and tools, the corpus has been aligned and segmented into phonemes and syllables using SPPAS (Bigi & Hirst 2012). A total of 20 files were manually annotated in prominences and semi-automatically segmented into major prosodic units using the Anamor software (<http://www.lattice.cnrs.fr/ressources/logiciels/anamor/>) (Avanzi et al. 2008). In addition, SLAM+, a tool that generates intonative contours automatically (<https://github.com/vieenrose/SLAMplus>) was developed to process intonative contours, particularly syntactic units, on a large scale (Liu et al. 2019). There arose here an interoperability problem and a tool was developed as part of the project, to retrieve under PRAAT the data encoded in formats dedicated exclusively to syntax processing (Arborator and CONLL). Concerning the functional analysis, a first set of data was used to make hypotheses on the phonetic marking of the focus in Naija (Simard et al. 2019). Another has been processed since July 2019 to study the intonative contours of pre-kernels and macrosyntactic nuclei. This new step should allow us to answer a set of questions related to the encoding of the informational structure of the message in Naija, including: in the initial position of statement, are there specific intonative markers of the pre-kernels and are the observed intonative variations correlated to the informational status of the element included in the pre-kernel (topic

vs. frame; topic active, vs. accessible, introduced, reactivated, etc.) These studies on the intonational markers of framing, topicalization and focusing operations, are still ongoing.

4.2.3. Morphosyntactic analysis

The main objective is to tag, gloss, and parse the 500 Kw corpus, using state of the art NLP tools. In the process, a 100 Kw gold-standard, manually corrected treebank (NTB), has been produced. A new parser, using word embedding and neuronal technology will soon be trained on these files and used to annotate the remaining part of the corpus (Reference Naija Corpus and Deuber Corpus). The evaluation of the automatically annotated corpus (NRC) will start early 2020. Finally, the remaining 400 Kw data will be analysed by the parser trained on the gold-standard treebank, providing the 500 Kw treebank to be compared with ICE-Nigeria and more generally available for use by the sociolinguistic work package.

Glossing and POS tagging. To start the annotation process, a first sample text was tagged with a model trained on English. Insofar as most of the lexicon of Naija is borrowed from English, and its meaning is transparent, the glossing was kept to a minimum. Function words do not have glosses beyond their morphological features, and only Naija lexical innovations were glossed (e.g. *pikin* ‘child’, *patapata* ‘full’). The POS annotation was manually corrected and a first dictionary of the function words and most common lexical items of Naija was created, containing the form, some orthographic variants, the POS tag, and an English gloss if necessary. This dictionary was then used on a dozen text samples inside the Elan-Corpa tool (Chanard 2014), an extended version of the Elan tool3 (Sloetjes 2014), which proposes the dictionary’s POS for each token for validation by the annotator. Through this semi-automatic process, the dictionary was enriched and later on used by the automatic tagger that was developed for the project. The POS tags follow the UD conventions (Nivre et al. 2016) with the caveat that some changes were made to accommodate the specificities of the Naija system. For example, Naija has three copulas, *be*, *dey* and *na*, among which two are tagged as VERB (and ‘be’) and one, also used as a focus particle, is tagged as PART (‘it is’). Regularly, the POS tagger is trained again on the corrected tags and thus improved in a bootstrapping loop.

Annotation guidelines. The annotation process for the samples was organized collectively, where each file was assigned to one of the three annotators. They were allowed to discuss the difficult cases among each other. At the end of this process, the annotation was consolidated through the use of a dictionary that was controlled independently and applied to the corpus. The final adjudication was done by an expert adjudicator on every single file. In this process some amendments had to be discussed more widely in the UD community. The annotators are asked to verify their annotations by means of an annotation guide and to report directly into the guide any decision that is not directly derived from it. We thus have an annotation guide that undergoes constant refinement.

From UD to SUD. A preliminary result of the syntactic annotation was published on the UD project website as a mini-corpus pilot of 4 files with a grammatical sketch⁶. After this publication, we made a paradigm shift in the annotation of the corpus. It was decided to develop an enriched syntax annotation for Naija (Syntactic Universal Dependency, SUD) (Gerdes et al. 2018) compatible with Universal Dependency’s “classical” model.

⁶ http://universaldependencies.org/treebanks/pcm_nsc/index.html.

5. Half-way assessment

5.1. Planned results

The NSC project plans to produce a 100 Kw gold standard treebank for Naija (manual correction); a 400 Kw treebank for Naija (automatic annotation); syntactic annotation guidelines for Naija; a tagger and a glosser for Naija; a dependency parser for Naija (MATE trained on our gold standard treebank); a 500 Kw treebank for Nigerian English (ICE Nigeria analysed with the English Stanford parser). The whole process is expected to deliver the annotated corpora for analysis in early 2020.

5.2. Corpus editing and annotation

The early stages of the NSC corpus construction (fieldwork, alignment, transcription, translation) went quite smoothly and ran ahead of schedule. The corpus construction was completed in December 2017 in Nigeria by the annotation team at the University of Ibadan, under the supervision of the Principal Investigator. In the 10 survey points, more than 380 files were initially edited (representing 31 hours of recording) with corresponding metadata (approx. 350 speakers). After reviewing the audio quality and contents, the corpus was sized down to 321 files and 343 speakers, constituting the NSC reference corpus. By mid-2019, the size of our corpus is as follows (Table 1):

Golden Treebank	Reference Corpus	Deuber (2005)
80 files	241 files	50 files
96 kw	302 kw	100 kw
7 hours	24 hours	10 hours

Table 1. NSC Corpus

Inconsistencies in the morphosyntactic annotation forced us to run systematic tests and extra manual corrections on the Golden Corpus, which have entailed some delays. Then, in order to improve the quality of the automatic parsing, an extra revision of the macro punctuation and sentence alignment was done for the remaining 241 files of the Reference Corpus.

5.3. General problems faced by the project

5.3.1. Sampling

If the geographical sampling is acceptable, the result is not balanced in terms of sex (women represent only 1/3 of the sample; cf. Figure 6), age (Figure 7) and education (Figure 8).

Very few young and elderly speakers were recorded and more than 50% of the speakers were highly educated: more than 50% were graduates and more than 20% were higher education students.

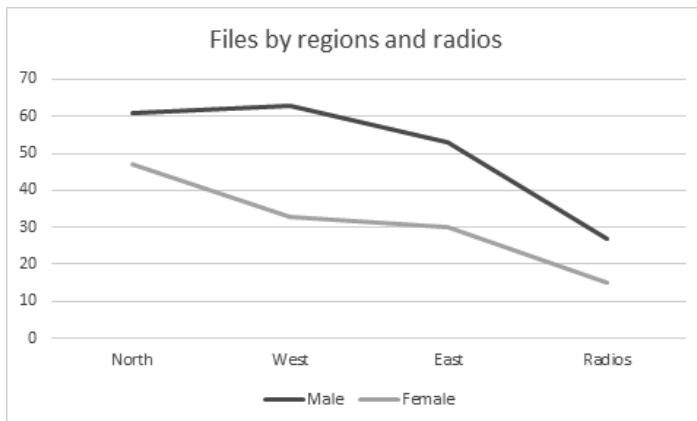


Figure 6

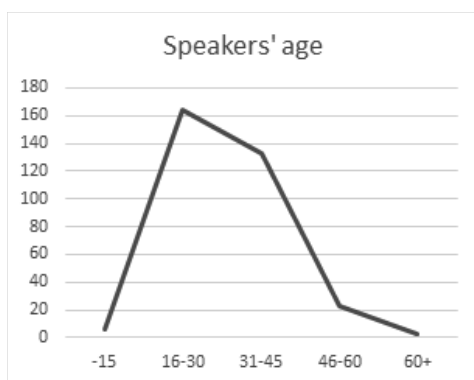


Figure 7

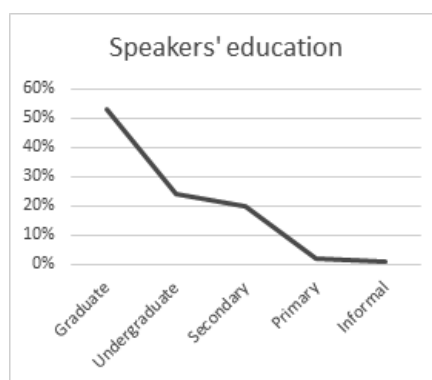


Figure 8

This paints a picture of Naija as a language proudly spoken by highly educated people, who codeswitch confidently from Naija to English, far from the usually prevalent picture of a limited code (a pidgin) used by illiterates who are linguistically impaired.

As a conclusion, we can state that the corpus is expected to give a good representation of the language; not an exhaustive representation of its status or its speakers. However, it contradicts the widely spread prejudice that Naija is a language associated with illiterates, and a danger for Nigerian education and English in Nigeria.

5.3.2. Transcription

Orthography: due to unstable orthography, annotators have not been consistent. A lemma was selected in the tagging process, and variants were kept in the transcription, e.g. ‘thing’: *thing, ting, tin* (lemma: *ting*); ‘their’: *their, deir, dier* (lemma: *deir*); ‘there’: *there, dere, dier* (lemma: *dere*); ‘him’: *him, im, in* (lemma: *im*)

Phonetics: The transcription of phonetic variants has not been consistent, which means that this aspect of variation cannot be studied in the corpus, e.g. ‘make’: [*mek, me, ma, mo*]; ‘him’: [*him, im, i*]; ‘them’: [*dem, dem, de*].

Morphology: the speakers' habits, some orthographic decisions have been taken to eliminate lexical ambiguities that have strong morphological and syntactic implications. When annotators did not respect these decisions, the variants were kept in the text and the NSC orthography indicated as a lemma. This is the case with morphemes derived from serial verb constructions, in which the second verb has been grammaticalised, e.g. the complementizer which we have chosen to write *sey* instead of *say* (as it is usually spelt by Naija speakers) to disambiguate it from the verb *say*. Likewise, we have chosen to write the future auxiliary *con* instead of *come*.

As a first contribution to the establishment of a standard orthography of Naija, it has been decided to publish online a normalised version of the corpus aimed at a wider Nigerian audience.

5.3.3. Morphosyntactic analysis

Decisions had to be taken concerning the tagging of various items, e.g. the copulas *be*, *dey* and *na*, which can be tagged as auxiliaries following the UD guidelines, verbs or particles. Since *na* (a focus marker which can be also used as a predicative copula) cannot be combined with TAM markers or negation, we have analysed it as particle, contrary to *be* and *dey*, which have been analysed as verbs.

Another problem concerns a category that has been labelled “property items” by Mazzoli (2013) in a bid to avoid calling them either verbs or adjectives. Faraclas clearly argues for the inexistence of adjectives in Nigerian Pidgin: “*there is no category ‘adjective’ in Nigerian Pidgin. Most of the items which convey the same meanings as do adjectives in other languages are stative verbs in NP. Stative verbs take the same arguments and modifiers in the same combinations and the same order as do other verbs.*” (Faraclas 1989: 132)

Indeed, our Naija corpus corroborates Faraclas’s analysis of Nigerian Pidgin, in that numerous examples of “property items” are used as stative verbs, both intransitive, as in (3) and (4) and transitive as in (5):

3. *Women sef, we bad.* ‘We, women, we [are] bad.’ [P_JOS_14]
4. *Di meat sweet o!* ‘The meat [is] quite tasty!’ [P_JOS_20]
and transitive, as in (5)
5. *Di weather dey sweet us.* ‘We enjoy the weather (lit. the weather [is] nice to us)’. [P_JOS_20]

They combine with TAM auxiliaries, e.g. the Future AUX *go* in (6) and the Imperfective *dey* in (8):

6. *Di fruit go sweet.* ‘The fruit will [be] nice.’ [P_IBA_31]

They combine with the negative particle *no*, as in (7):

7. *Belle no sweet am at all* ‘He is not happy at all. (lit. the stomach does not satisfy him at all.)’ [P_WAZK_07]

In (8), the term *sweet* is part of the comparative serial verb construction *sweet ... pass*, lit. ‘be sweeter’:

8. *Na dat one dey sweet me pass.* ‘It’s that one I like best. (*lit.* It’s that one that satisfies me most.)’ [P_IBA_02]

This leads Faraclas to analyze nouns modified by property items (e.g. big money, small work, bad name) as relative clause constructions: “*Since the category ‘adjective’ does not exist in NP [...] and because of the fact that the only type of clause in the language which may serve to modify nominal elements is the relative clause, the label ‘adjective clause’ is not employed here, ‘relative clause’ being used instead.*” (Faraclas 1989: 75)

However, if terms like *sweet, smooth*, etc. function mainly as a verbs in our Naija corpus, others like *next, last, waye* ‘dubious’, etc. only function as adjectives.

In conclusion, in our Naija corpus, there is no clear case in favor of choosing either adjective (ADJ) or verb (VERB) as a POS tag for those items. For comparative purposes, we have decided to tag as ADJ items that can function both as adjectives and verbs, while we keep the tag VERB for those that can only function as verbs. Likewise, we tag NOUN items that can function both as nouns and verbs.

5.4. Exploratory results of the corpus-based analysis

Several papers have been published exploring some striking properties of Naija, as they already appear in the corpus.

(Čéplö & Manfredi 2019) have made a first attempt at assessing morpho-syntactic variation in Naija. (Simard et al. 2019) make a first presentation of the prosody of the language. They deal with focus and prominence types and more specifically, with the prosodic encoding of narrow focus. They were able to show that narrow focus (*na+* focused element construction) is conveyed mainly with duration prominence (See fig. X)

9. *De fit say de dey strike.* ‘they can declare industrial action’ [P_IBA_21]

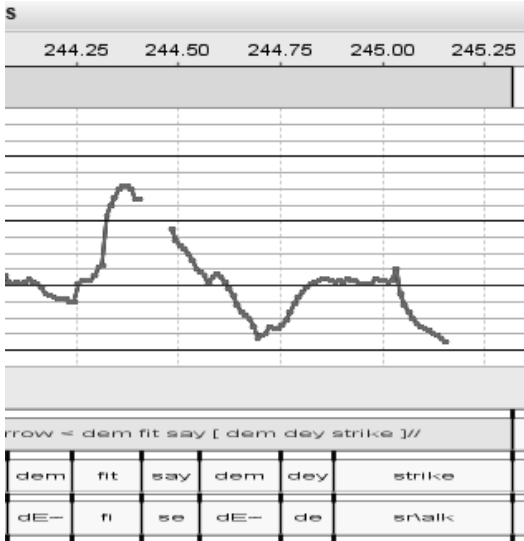


Figure 9. Pitch track of Example 9, illustrating the longer duration on the focused element, strike.

Several papers deal with the syntax of Naija, and, among other things, with serial verb constructions (Caron et al. 2019) and clefts (Caron 2020). The basic element in the structure of clefts in Naija is the focus particle *na*, as in (6):

10. *na nineteen eighty four >+ wey de born me //* ‘it’s in nineteen eighty-four that I was born.’ [P_KAD_09]

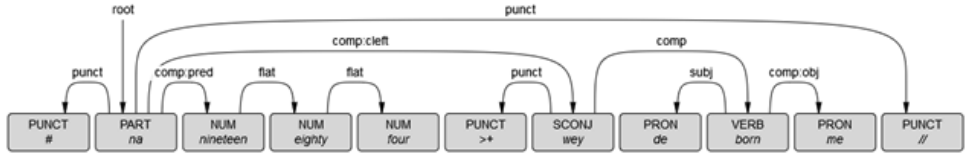


Figure 10. Dependency structure of Example 10.

Clefts in Naija show a great variety of structures and clear innovations from Nigerian Pidgin. Naija clefts have four variants, illustrated in Table 2: *wey*-clefts (a), with a relative clause introduced by the relativizer *wey*; bare clefts (b), where the relativizer is omitted, resulting in a bare relative clause; zero clefts (c) where both the copula and the relativizer are omitted; and double clefts (d), where the relativizer *wey* is replaced by a repetition of the copula *na* followed by an expletive invariable 3sg pronoun: *im*.

a	<i>wey</i> -cleft	<i>na 1984 wey de born me</i>	‘(It’s) in 1984 (that) I was born.’
b	bare cleft	<i>na 1984 Ø de born me</i>	
c	zero cleft	<i>1984 de born me</i>	
d	double cleft	<i>na 1984 na im de born me</i>	

Table 2. The four structures of Naija clefts

We have quantified the relative use of these structures in Naija in a sub-section of 9621 sentences (almost 150 000 tokens) that constitute the syntactic treebank mirroring the social and geographic sampling of the full corpus, and compared those figures with Faraclas (2013), a presentation of the structures of Nigerian Pidgin with good data analysis. Using our own terminology, Faraclas’s figures highlight 3 main patterns representing fairly evenly cleft constructions in NP: *wey*-clefts (41%); bare clefts (39%) and zero-copula clefts (17%). Our own figures are respectively 1%, 89%, and 1%, with the rest of cleft patterns taken up by double clefts (9%). This shows a tendency in Naija, over the past 30 years, to marginalize *wey*- and zero-copula clefts, in favor of bare clefts, and give birth to a new pattern absent in Faraclas’s description, called double cleft, which seems to replace *wey*-clefts. In the double cleft construction, an emerging relative pronoun (*na im* → [nāĩ/nā] ‘who, which’) which is used only in this construction, replaces the relativizer *wey*, which is becoming specialized in modifying relative clauses.

This shows that Naija is changing fast while the Nigerian Pidgin (NP) described in (Faraclas 2013) has not changed much from the one described in (Faraclas 1989), based on a large corpus gathered by the author in Rivers State thirty years ago. However, there is a stunning difference between NP, a creole spoken in the Niger delta area, and the Naija documented in our corpus, a pidgincreole spoken as L2 by a vast majority of Nigerian.

6. Conclusion

To conclude this half-way assessment of the NaijaSynCor project, it should be noticed that it has already produced a large fine-grained annotated corpus, the first for creoles and pidgins studies⁷. This corpus will serve as a basis for the development of Naija, through the publication of a large corpus in stabilized orthography, a dictionary, and a grammar.

The project has exceeded the time allotted for annotation, which leaves a limited time for analysis. However, we anticipate that the quality of annotation and the powerful tools currently developed or improved by the research team (Trameur, Grew) will help us compensate for the time limitation. 2020 will be devoted to the exploration and evaluation of the corpus, and we intend to finish the NaijaSynCor research project by the end of the year.

Last but not least, the NSC project has developed and improved a series of NLP tools that can be extended and adapted to the study of lesser-described African languages. However, since the efficiency of the new generation of NLP tools relies more and more on access to very large quantities of raw data, the limited resources available for minority languages will impact the quality of the results. The output of such tools will probably be used as annotation propositions that will have to be manually validated by researchers. This nevertheless opens the way to a new methodology for the documentation, description and development of minority languages.

Abbreviations

ADJ	Adjective	POS	Part of Speech
AUX	Auxiliary	TAM	Tense-Aspect-Mood
NP	Nigerian Pidgin	UD	Universal Dependencies
NSC	NaijaSynCor	VERB	Verb

References

- Agheyisi, Rebecca Nogieru. 1984. "Linguistic implications of the changing role of Nigerian Pidgin English". *English World-Wide* 5, 211–233.
- Ajibade, Yetunde A., Beatrice Bunmi Adeyemi & Emmanuel Olajide Awopetu. 2012. "Unity in Diversity: The Nigerian Youth, Nigerian Pidgin English and the Nigerian Language Policy". *Journal of Educational and Social Research* 2(3), 289-295.
- Andersen, Hanne Leth & Henning Nølke (eds.). 2002. *Macro-syntaxe et macro-sémantique: actes du colloque international d'Århus, 17-19 mai 2001*. Berne: Peter Lang.
- Avanzi, Mathieu, Anne Lacheret & Bernard Victorri. 2008. Analor, un outil d'aide pour la modélisation de l'interface prosodie-grammaire. CERLICO, 27-46. <https://halshs.archives-ouvertes.fr/halshs-00636544>.
- Ayafor, Miriam & Melanie Green. 2017. *Cameroon Pidgin English: a comprehensive grammar*. Amsterdam: John Benjamins.
- Bakker, Peter. 2008. "Pidgins versus Creoles and Pidgincreoles". In: Silvia Kouwenberg & John Victor Singler (eds.). *The handbook of Pidgin and Creole studies* [Blackwell Handbooks in Linguistics]. Chichester, West Sussex/Malden, MA: Wiley-Blackwell Pub, 130–157.

⁷ See however a corpus of Cameroon Pidgin (Ozón et al. 2017), which has been automatically tagged for POS and used as a basis for a grammar of the language (Ayafor & Green 2017).

- Berrendonner, Alain. 1990. "Pour une macro-syntaxe". *Travaux Linguistiques de Gand* (21), 25–36.
- Bigi, Brigitte, Bernard Caron & Abiola S. Oyelere. 2017. "Developing Resources for Automated Speech Processing of the African Language Nijja (Nigerian Pidgin)". *8th Language and Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics*. Poznań, Poland. <https://hal.archives-ouvertes.fr/hal-01705707>, 441–445.
- Bigi, Brigitte & Daniel Hirst. 2012. "SPeech Phonetization Alignment and Syllabification (SPPAS): a tool for the automatic analysis of speech prosody". *Speech Prosody*, Shanghai (China): Tongji University Press, 19–22.
- Blanche-Benveniste, Claire, Mireille Bilger, Christine Rouget, Karel van den Eynde & Piet Mertens. 1990. *Le français parlé: études grammaticales*. Paris: CNRS.
- Boersma, Paul & David Weenink. 2013. *PRAAT: Doing phonetics by computer*. <http://www.fon.hum.uva.nl/praat/> (30 November, 2013).
- Caron, Bernard. 2020. "Clefts in Nijja, a Nigerian pidgincreole". *Linguistic Discovery* 17(1). 149–174.
- Caron, Bernard, Marine Courtin, Kim Gerdes & Sylvain Kahane. 2019. "A Surface-Syntactic UD Treebank for Nijja". In: *Proceedings of the 18th International Workshop on Treebanks and Linguistic Theories (TLT, SyntaxFest 2019)*. Paris, France: Association for Computational Linguistics, 13–24. <https://doi.org/10.18653/v1/W19-7803>.
- Čeplö, Slavomír & Stefano Manfredi. 2019. Assessing morpho-syntactic variation in Nijja (Nigerian Pidgin): a corpus-driven study. Presented at the 2019 SPCL summer meeting, University of Lisbon, June 17-19, 2019, Lisbon, Portugal.
- Chanard, Christian. 2014. *ELAN-Corpa-V4.7.3*. http://lilac.vjf.cnrs.fr/res_ELAN-Corpa.php.
- Cresti, Emanuela. 2000. *Corpus di italiano parlato*. 2 vols. Firenze, Italie: Accademia della Crusca.
- Deuber, Dagmar. 2005. *Nigerian Pidgin in Lagos. Language contact, variation and change in an African urban setting*. London: Battlebridge Publications.
- Deulofeu, Jose, Kim Gerdes, Sylvain Kahane & Paola Pietrandrea. 2010. Depends on what the French say: Spoken corpus annotation with and beyond syntactic function, 1–8. <https://halshs.archives-ouvertes.fr/halshs-00649791> (1 December, 2015).
- Egbokhare, Francis O. 2004. "Language and politics in Nigeria". In: Kólá Owólábi & A. O Dasylyva (eds.). *Forms and functions of English and indigenous languages in Nigeria: a festschrift in honour of Ayo Banjo*. Ibadan: Group Publishers, 507–22.
- Elugbe, Ben Ohiomamhe & Augusta Phil Omamor. 1991. *Nigerian Pidgin: (background and prospects)*. Ibadan: Heinemann Educational Books Nigeria PLC.
- Faraclas, Nicholas. 1996. *Nigerian Pidgin*. London ; New York: Routledge.
- Faraclas, Nicholas. 2013. Survey chapter: "Nigerian Pidgin". In: Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.). *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://apics-online.info/surveys/17>.
- Faraclas, Nicholas Gregory. 1989. *A grammar of Nigerian Pidgin*. University of California at Berkeley PhD.
- Fleury, Serge & Maria Zimina. 2014. "Trameur: A Framework for Annotated Text Corpora Exploration". In: Tsujii, Junichi & Jan Hajic (eds.). *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: System Demonstrations. August 2014, Dublin, Ireland*. Dublin City University and Association for Computational Linguistics, 57–61.
- Gerdes, Kim, Bruno Guillaume, Sylvain Kahane & Guy Perrier. 2018. "SUD or Surface-Syntactic Universal Dependencies: An annotation scheme near-isomorphic to UD". *Universal Dependencies Workshop 2018*. Brussels, Belgium. <https://hal.inria.fr/hal-01930614>.
- Guillaume, Bruno, Guillaume Bonfante, Paul Masson, Mathieu Morey & Guy Perrier. 2012. "Grew : un outil de réécriture de graphes pour le TAL". <https://hal.inria.fr/hal-00760637> (17 November, 2019).
- Gut, Ulrike. 2014. "ICE Nigeria". *SourceForge*. <http://sourceforge.net/projects/ice-nigeria/> (6 August, 2014).
- Huber, Magnus. 1999. *Ghanaian Pidgin English in Its West African Context: A Sociohistorical and Structural Analysis*. John Benjamins Publishing.
- Kahane, Sylvain & Paola Pietrandrea. 2012. *La typologie des entassements en français*. Vol. 1. SHS Web of Conferences. <http://www.shs-conferences.org/>.

- Lewis, M. Paul, Gary F. Simons & Charles D. Fennig. 2013. Nigeria. *Ethnologue: Languages of the World, Seventeenth edition*. Dallas, Tex.: SIL International. (22 February, 2014).
- Liu, Luigi (Yu-Chen), Lacheret-Dujour & Nicolas Obin. 2019. "Automatic modelling and labelling of speech prosody: what's new with SLAM+?". *International Congress of Phonetic Sciences (ICPhS)*. Melbourne, Australia. <https://hal.sorbonne-universite.fr/hal-02119926>.
- Mazzoli, Maria. 2013. *Copulas in Nigerian Pidgin*. Pavia: University of Pavia, PhD dissertation.
- Nivre, Joakim, Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Jan Hajic, Christopher D. Manning, Ryan McDonald, et al. 2016. "Universal Dependencies v1: A Multilingual Treebank Collection". In: Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Helene Mazo, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*. Paris, France: European Language Resources Association (ELRA).
- Ozón, Gabriel, Miriam Ayafor, Melanie Green & Sarah Fitzgerald. 2017. "The spoken corpus of Cameroon Pidgin English". *World Englishes* 36(3), 427–447. doi:10.1111/weng.12280.
- Rickford, John R. 1987. *Dimensions of a Creole continuum: history, texts & linguistic analysis of Guyanese Creole*. Stanford, Calif.: Stanford University Press.
- Simard, Candide, Anne Lacheret-Dujour & 'Biola S. Oyelere. 2019. "Broad and narrow focus marking in Naija (Nigerian Pidgin): the role of prosody". In: Sasha Calhoun, Marija Tabain Escudero & Paul Warren (eds.). *Proceedings of the 19th International Congress of Phonetic Sciences*. Canberra, Australia: Australasian Speech Science and Technology Association Inc. http://intro2psycholing.net/ICPhS/papers/ICPhS_3956.pdf, 3907–3911.
- Simon, Anne Catherine & Liesbeth Degand. 2011. "L'analyse en unités discursives de base : pourquoi et comment?". *Langue française* 170(2), 45–59.
- Sloetjes, Han. 2014. "ELAN: Multimedia Annotation Application". In: Jacques Durand, Ulrike Gut & Gjert Kristoffersen (eds.). *The Oxford handbook of corpus phonology*. Oxford: Oxford University Press 305–320.
- Tagliamonte, Sali. 2012. *Variationist sociolinguistics: change, observation, interpretation* [Language in Society 40]. Malden, MA: Wiley-Blackwell.

Part II



Structural properties of languages and their theoretical frames

■ Clause chaining in Bambara

Abstract

Bambara (Manding < West Mande < Mande < Niger-Congo) has a specialized clause chaining (cosubordinative) construction where the first clause has a verb full-fledged for TAM and polarity, and verbs of subsequent clauses are in infinitive; non-initial clauses are not embedded into the initial one and cannot be therefore regarded as subordinate. Characteristics of the Bambara clause chaining are analyzed: scope of operators (illocutionary force, negation), TAM characteristics, same- and different-subjectness. From the evidence of peripheral Manding varieties, a hypothesis concerning the origin of infinitive from an archaic preposition is advanced.

Keywords: clause chaining, cosubordination, infinitive, Bambara, Manding

1. Introduction

Clause chaining, as a clause linking type distinct from both coordination and subordination, for the first time was described by Olson (Olson 1981), then by Foley & Van Valin (Foley & Van Valin Jr. 1984). Another term for this linking type is “cosubordination”. Its main characteristics are:

- in a sequence of clauses, one of them (either initial or final, depending on language) is morphologically full-fledged (its verb is inflected as a finite), and the others are morphologically marked as dependent: most often, their verbs appear in forms which cannot be used in independent clauses (converbs, infinitives, participles, medial verbs);
- the scope of the inflectional category or operator (mainly TAM) marked normally on the “quasi-head” clause, is the entire clause chain, i.e. both “quasi-head” and “quasi-dependent” clauses;
- the clauses whose verbs lack full-fledged inflection (and therefore appear as “morphologically dependent”) are not embedded into the “quasi-head” clause, but rather co-ranked with it.

Most European languages have no clause chaining; clauses expressing sequential events are simply coordinated.

Depending on the position of the “quasi-head” clause in the clause chain, two strategies can be realized:

- 1) the first verb appears in a finite form (the first clause is “quasi-head”), the verbs of the subsequent clauses are specially marked (posterior clause chaining). This

strategy is widely spread in African languages, and rather rare elsewhere (Longacre 2007: 417);

- 2) the last verb appears in a finite form (the last clause is “quasi-head”), the preceding verbs are specially marked (anterior clause chaining). This type is widespread in Asia, New Guinea, in Caucasus, in Ethiosemitic languages.

It is often mentioned that “anterior chaining is typologically associated with OV basic order, and posterior chaining is associated with VO order” (Haspelmath 1995: 23), see also (Longacre 2007: 417).

In a more recent paper, Foley (2010: 40) re-analyzes, on the syntactic ground, the cosubordination as a variety of coordination: “in such structures the verbal inflectional I features of the verb of the last clause (i.e., the full-fledged clause – *IV*) typically have scope over the preceding medial or dependent clauses. Verbs in medial clauses are commonly stripped down inflectionally in comparison to final verbs, as a reflection of this scope dependency. But, in fact, the inflectional I categories of the verb of the final clause do not belong to it, but rather to the structure as a whole”. Certainly, Foley is right with respect to the syntactic level: from the very beginning, it was evident that cosubordination was close to coordination syntactically, and resembled subordination morphologically. It seems expedient to maintain the term “cosubordination” exactly to highlight the ambiguous nature of this construction.

An important direction of study of clause linking (and, in particular, cosubordination) is the scope of operators. As shown in Bickel (2010), the scope of different operators (illocutory force, negation, tense) may vary considerably both among languages and within one language. It means that cosubordination is a complex and variable phenomenon, and its characteristics in each language should be studied in detail.

In this paper, the clause chaining construction¹ in Bambara will be analyzed. In section 2, general information about Bambara and, in particular, about Bambara infinitive, is provided. In section 3, a general characteristics of the Bambara clause chaining construction is given. In sections 4, 5, and 6, the main characteristics of the clause chaining construction are analyzed in more detail. In section 7, a hypothesis concerning the diachronic origin of the Bambara infinitive is advanced.

2. Introductory information about Bambara

2.1. Some generalities

Bambara (Manding < Western Mande < Mande) is spoken mainly in Mali by some 4 million L1 speakers and by further 10 to 12 million L2 speakers. By default, by “Bambara” is meant the so-called Standard Bambara, a variety based on the dialect of Bamako (which is influenced by closely related Maninka).

Bambara is a tonal language (two level tones whose attribution is regulated by a set of rules, see among others (Vydrin 2016; Vydrin 2019: 31–36).²

¹ The terms “clause chaining” and “cosubordinative construction” will be used in this paper as more or less synonymous.

² In the Bambara examples, tone marking and glossing follows the principles accepted for the Bambara Reference Corpus (Vydrin et al. 2011).

As all Manding languages, Bambara is characterized by a rigid word order. In the verbal clause, it is as follows: S Aux (O) V X, where S stands for subject; Aux is a non-verbal auxiliary word encoding the TAM and polarity meanings;³ O is direct object (obligatory for a transitive verbal construction; the absence of O makes a construction intransitive); V is a verbal predicate; X is an oblique, i.e. an adverbial or postpositional phrase.

- (1a) *Mùso` bε jéε` mìnε kò` lá.*
 woman\ART IPFV fish\ART catch brook\ART in
 ‘The/a woman catches fish in a brook’.

Lability is highly typical for Bambara verbs. Very productive is P-lability, both decausative and passive (1b), the latter being absolutely regular; A-lability is rather marginal (Vydrin 2019: 201–214).

- (1b) *Jéε` bε mìnε kò` lá (mùso` fε).*
 fish\ART IPFV catch brook\ART in woman\ART by
 ‘Fish is caught in a brook (by the/a woman)’.

2.2. Infinitive

Infinitive plays a central role in the Bambara clause chaining construction and needs therefore a more detailed presentation. Infinitive constructions have been subject of several studies (Dumestre 2003: 384–396; Dombrowsky-Hahn 2012; Vydrin 2014; Vydrin 2019: 409–430); see also Creissels (2009: 73–74) about infinitive in Kita Maninka (which does not differ from the Bambara infinitive). Here follow some basic data on their functions and peculiarities.

Infinitive in Bambara is an analytical construction, its marker is an auxiliary word *kà*, which occupies the same slot as auxiliaries (“predicative markers”) in finite verbal constructions. As it is typical of infinitives cross-linguistically, in the Bambara infinitive construction the subject cannot be formally expressed, the zero subject being normally co-referent with the subject or the direct object of the matrix verb (see section 6 about more complicated cases). As for the other arguments and obliques, they are presented in the infinitive construction exactly as with a finite verb.

An infinitive construction cannot be used in the function of a subject or direct object; on the other hand, it can appear in the position of topic and be coreferent with a subject (2) or object (3) pronoun.

- (2) *[Kà` dúnan-w fòroba-εn]ᵣ òᵢ kón-nen dòn.*
 INF foreigner-PL collectively-chase that prohibit-PTCP.RES ID
 ‘The mass expulsion of non-nationals shall be prohibited’ (lit. ‘To chase foreigners collectively, this is prohibited’) [Afiriki ka hadamaden].

³ In the Mandeist linguistic tradition, these auxiliaries are usually referred to as “predicative markers”. This term will be also used in the present paper, as a synonym of “auxiliary”.

- (3) [Kà jéǵe` món]_p án tɛ ò_i ké.
 INF fish\ART fish 1PL IPFV.NEG that do
 ‘Fishing, we do not do it’.

An infinitive construction coreferent to the direct object (expressed by a pronoun) of the matrix verb is not necessarily topicalized, it may also follow the matrix verb (4).

- (4) ... ù bɛ à_i fǒ [kà jí goni` dǒnin sìǵi_i
 3PL IPFV 3SG say INF water hot\ART a.little sit
 í dá` kónǒ]...
 2SG mouth\ART inside
 ‘... they say to pour some hot water into your mouth...’ [Avenement de Da].

The first most prototypical function of the Bambara infinitive is that of a predicate of an argument clause or an adverbial clause.⁴

Matrix verbs with a valency on infinitive can be subdivided in Bambara into three groups: modal, aspectual, and manner verbs.

Modal verbs can be intransitive or reflexive (*bàn* ‘refuse’, *dése* ‘fail’, *jǵà* ‘try, strive’, *jùne* ‘forget’, *sé* ‘be able’, *sòn* ‘agree’, *túgu* ‘do on purpose’, and some others), in which case the zero subject of the infinitive is necessarily coreferent to the subject of the matrix verb (5a), or transitive (*dème* ‘help’, *yàmaruya* ‘allow’, *bàli* ‘hinder’, and some others), the zero subject of infinitive being coreferent with the direct object of the matrix verb (6).

- (5a) à ka kóǵri`_i bɛ sé ø_i kà móbili` fá.
 3SG POSS cotton\ART IPFV reach INF car\ART fill
 ‘His cotton can fill a car’ [Baara kalan ka nyésin].

- (6) Ñká ò tɛ án_i bàli ø_i kà à jéǵini.
 but that IPFV.NEG 1PL hinder INF 3SG solve
 ‘But this does not prevent us from solving it’ [Balimamusoden].

No pause can be inserted between the matrix verb and the infinitive marker;⁵ if the matrix verb is negated, the scope of negation includes the infinitive too (5b).

- (5b) à ka kóǵri` tɛ sé kà móbili` fá.
 3SG POSS cotton\ART IPFV.NEG reach INF car\ART fill
 ‘His cotton cannot fill a car’.

⁴ The second function, that of the predicate of a non-initial clause in a cosubordinative construction, is the main topic of this paper; it will be considered in the subsequent sections.

⁵ According to Dumestre (2003: 390), in constructions with modal and aspectual verbs, the matrix verb “cannot be separated from *kà* by any element or any suspensive pause”; cf. the same opinion of Dombrowsky-Hahn (2012: 54). Contrary to this opinion, my informants allow insertion of adverbs after the matrix verb, and numerous examples of such insertions can be found in the Bambara Reference Corpus, e.g.: *án bɛ sé túgun kà dǒ fàra jíw háke` kán* ‘we can add some water again’ [Kibaru 140].

For these reasons, a construction of a modal verb with an infinitive can be probably regarded as monoclausal.

Aspect verbs (*kàn* ‘do earlier’, *dèli* ‘do at least once’, *fàma* ‘fail to do since long time’, *sòli* ‘do early in the morning’, *tìlen* ‘do during the daytime’, etc.) are similar with the modal verbs in many respects, but there is at least one important difference: when negated, the scope of negation does not extend to the infinitive construction (7b).

(7a) *nìn dón, Waawere sòli-la kà bó n' à ka npálan` yé...*
 this day Waawere do.early-PFV.INTR INF exit and 3SG POSS bag PP
 ‘That day Waawere left early in the morning with his bag...’ [Jekabaara 142].

(7b) *Á ma sòli kà bó à ka bùgu` kónɔ.*
 3SG PFV.NEG do.early INF exit 3SG POSS hut\ART in
 ‘He did not leave his hut early’ (i.e., he left it, but it was not early) [Kolonkise 10].

“Manner verbs” described by Dombrowsky-Hahn (2012) express rate (*bòli* ‘run’, *táama* ‘walk’, *sùulusáala* ‘go very slowly’), intensity (*gírín* ‘do suddenly, sharply’, *bála* ‘do sharply’), or other aspects (*móntɔ* ‘bend’, *ɣúnuma* ‘crawl’) of a motion. In fact, the same verbs (or most of these verbs) can appear as modifiers not only for motion verbs, but for other semantic types of verbs as well.

(8) *Nègeso dílanna` gírín-na kà à jìninka.*
 bicycle repairer\ART rush-PFV.INTR INF 3SG ask
 ‘Repairer of bikes asked him promptly’ [Jekabaara 21].

As shown in (Dombrowsky-Hahn 2012: 57–58), the scope of negation of manner verbs varies: it may extend to the manner verb only, or to the second verb (the infinitive), or to both.

If an infinitive does appear in the function of the predicate of an adverbial clause, it can be introduced without conjunction, in which case it expresses the meaning of purpose (9).

(9) ... *à ma mùso sòrɔ kà à fúru Ségukɔrɔ yàn...*
 3SG PFV.NEG woman find INF 3SG marry Segukoro here
 ‘... he has not found a woman to marry here, in Segukoro’ [Chroniques amoureuses].

However, most often a purpose infinitive is introduced by a conjunction *wálasa* ‘in order to’ (10). There are also some other conjunctions capable to introduce infinitive clauses: *sáni* ‘before; instead of’, *fó* ‘until’, *kó* ‘that’, *jànkó ~ sànkó* ‘the more so’.

(10) *Án ye táama jan in ké wálasa kà í yé.*
 1PL PFV.TR walk long this do in.order.to INF 2SG see
 ‘We have overtaken this long journey to see you’ [Sunjata ka maana].

3. Main characteristics of the Bambara clause chaining construction

In Standard Bambara, there is no specialized construction or form for the “dependent” clauses of the clause chaining, this function is assumed by the infinitive construction. It is the second main function of the Bambara infinitive which may be regarded as secondary with respect to the prototypical infinitive function represented in 2.2; however, it predominates quantitatively. In certain genres of Bambara texts, up to 80% of all the occurrences of infinitive fall on the cosubordinative constructions. As it was mentioned in the introductory section, the use of prototypically non-finite forms in quasi-dependent clauses is typical for clause chaining, therefore, Bambara represents no anomaly from the typological viewpoint.

Bambara is a posterior clause chaining language. The first clause is full-fledged; this clause is obligatory (i.e., in Bambara, no clause chaining construction is possible without an initial clause with a full-fledged finite verbal construction).

Most often, the infinitive clauses are introduced without any conjunction, as in (11).

- (11) ... *ù nà-na kà mùru kura` jìni, kà siralan kura`*
 3SG come-PFV.INTR INF knife new-ART search INF broom new-ART
jìni, kà filen kama` jìni.
 search INF calabash new-ART search
 ‘They came, they looked for a new knife, (they looked for) a new broom, (they looked for) a new calabash’ [Bamakò sigicogoya].

However, such clauses can be also introduced by the coordinative conjunctions *ni* ‘and’ (12, 13) or *aní* ‘and’, *wàlà* ‘or’, *wàlimà* ‘or’ (14), *kélen* ‘or’. If a chain contains more than two clauses, a coordinative conjunction can introduce each non-initial clause (12) or only one of them (13).

- (12) *Né yèrê dòn-na, ni kà nà yèlen à kàn, ni k’*
 1SG.EMPH self enter-PFV.INTR and INF come rise 3SG on and INF
à jìni.
 3SG search
 ‘I came in myself, climbed on it and searched for it’ [Chroniques amoureuses].
- (13) *Bàsékù ye dùga`-w ké nê yé ni kà wúli*
 Baseku PFV.TR blessing\ART-PL do 1SG.EMPH PP and INF rise
kà nà nê bilasira.
 INF come 1SG.EMPH see.off
 ‘Baseku blessed me, he rose and saw me off’ [Chroniques amoureuses].

- (14) ... à tòn bε bée ké fɛn-tigi yé, wàlimà kà bée ké fàantan yé.
 3SG PST IPFV all do thing-owner PP or INF all do poor PP
 ‘... he would make everybody rich or he would do everybody poor’ [Kibaru 536].

Even if the use of coordinative conjunctions for clause chaining is relatively rare,⁶ its possibility is just another argument in favor of the “syntactically coordinative” nature of the cosubordinative construction in Bambara, contrasting with its “morphologically subordinateness”.

4. Tense/aspect and clause chaining

The clause chaining construction in Bambara (and in other Manding languages) is most often referred to as “consecutive” or “sequential” (Creissels 2006: vol. 2, pp. 186–189; Dumestre 2003: 385–386; Vydrin 2014; Vydrin 2019: 422–424). These terms based on the grammatical semantics of this construction are generally justified, because in the great majority of cases it is used for consecutive punctive actions or events. The verb of the initial clause of the construction is prototypically marked for the perfective aspect (11-14), which is specialized in punctive actions; in this case the sequentative semantics of the clause chaining construction is doubtless.

Less frequently, but by no means exceptionally, the verb of the initial clause may appear in practically all other TAM constructions existing in Bambara: imperfective (15), progressive (16), future (17), subjunctive (18, 19), conditional (20). Even in those cases, it can be still said that the actions expressed by the verbs of the clause chain are usually sequential.

- (15) ... sísàn sùngurun` dɔ-w, ò bε dòn móbili` lá kà tága
 now girl\ART some-PL that IPFV enter car\ART in INF go
 jàmana wére lá kà dɔn` ké yèn...
 country other in INF dance\ART do there
 ‘Nowadays, some girls take cars, go to other countries and dance there...’ [Dɔnkε-sunguru].

- (16) Á tága-ra à sɔrɔ sògo-w bé kà jógɔn fàga kà jógɔn tóbi.
 3SG go-PFV.INTR 3SG FIND animal-PL prog RECP kill INF RECP cook
 ‘She went on and met animals who were killing and cooking each other’ [Contes bambara 1979].

⁶ In the disambiguated subcorpus of the Bambara Reference Corpus (about 1.112.000 tokens, the state of June 1, 2019), the number of infinitive clauses is more than 28.600. If we assume that the clauses in clause chaining constructions constitute between 60 and 80% of the total number of infinitives, they score somewhere between 17.200 and 22.900. Of this number, the number of infinitive constructions introduced by *ni* equals 671; by *àní*, 261, and by *wàlimà*, 78, i.e. some 4 to 6%.

- (17) *Ò bènà í kùn` dí, kà dá-si` dí.*
 that FUT 2SG head\ART shave INF mouth-hair\ART shave
 ‘He will shave your head, he will shave your mustache’ [Maningances].
- (18) *Áw dén` bólo` ka cì, k’ ù tànga j̀̀ninsàn` mà.*
 2PL.EMPH child\ART hand\ART SBJV break INF 3PL protect smallpox ADR
 ‘Your child should be vaccinated and protected against smallpox’ [Bolociw].
- (19) *Ní pónpe kala` jí-bɔ-da` m̀n géren-na, án k̀nâ*
 if pump stick\ART water-exit-mouth\ART REL block-PFV.INTR 1PL PROH
ò dòn án dá k’ à f̀ye.
 that enter 1PL mouth INF 3SG blow
 ‘If the watering orifice of the pump gets blocked, let us not put it into the mouth and blow it out’ [Kibaru 41].
- (20) *Jí` mánà lánɔɔ kà sé háke` dó mà, í be*
 water\ART COND soil INF reach measure\ART some ADR 2SG IPFV
bàna-kise-w s̀rɔ à lá ...
 illness-grain-PL find 3SG in
 ‘If water soils to certain extent, you will find microbes in it...’ [Kibaru 547].

However, the sequential semantics of clause chaining constructions is rather a strong tendency than a strict rule. When the initial clause contains a TAM construction other than perfective, actions designated by the verbs of the subsequent clauses may be simultaneous,⁷ as in (21-23), otherwise, the first verb may express the mode of action designated by the second verb (24).

- (21) *Cèkɔɔba fila ninnu, ù bé kà kúma` lámɛn ni kà nàamù` lámìnɛ.*
 old.man two these 3PL PROG speech\ART listen and INF yes\ART keep
 ‘These two old men are listening to the speech and keep up the conversation’ [Maningances].
- (22) *Mɔɔgɔ k̀nâ búbaga` d̀gɔya kà í tó nt̀nkun` s̀anfê.*
 human PROH termite\ART humiliate INF REFL remain termite.hill\ART above
 ‘One does not insult termites while standing on the top of a termite hill’ [Sagesse bambara].

⁷ Contrary to what is said by Creissels (2009: 210) about the analogous construction in Kita Maninka, a Manding variety very close to Bambara. It is true though that the simultaneous meaning of the clause chaining construction in Bambara is not very frequent and subject to considerable lexical limitations.

- (23) *Cíkela` mánà d̀gukolo` s̀nɛ kà j̀nɛ ò wáleya ninnú kó,*
 farmer\ART COND land\ART cultivate INF forget that action these behind
à méen ô méen à nà j̀gi Ála lá gánsan.
 3SG continue DISTR continue 3SG CERT hope God in simply
 ‘If a farmer tills the land while forgetting about these actions, however long it may last, he will have to simply rely on God’ [Kibaru 546].

- (24) *Bàra` bé kà kòlonkolo kà nà.*
 gourd\ART PROG roll INF come
 ‘The gourd is coming rolling’ [Cíkela ni bara].

These examples, although not very frequent, prove that the consecutive meaning is not inherent in the clause chaining construction (even if a strong correlation is evident).

For this reason, I prefer the terms «clause chaining» and «cosubordination» for the construction in question. The terms «consecutive» or «sequential» fit well to characterize the predominant semantics of this construction, but less so as labels for the construction itself.

5. Scope of operators for the clause chaining construction

5.1. Illocutionary force: interrogation and imperative

A standard test for cosubordination is the scope of illocutionary operators: “a clause that is cosubordinate to a main clause obligatorily falls under the scope of illocutionary operators in the main clause, <...> apart from coordination, where the scope of such markers does not necessarily extend over both clauses, and also from subordination, where it is impossible to have conjunct scope» (Bickel 2010: 52).

In Bambara, general question is normally expressed by a particle *wà* whose position is at the end of a sentence. In a clause chaining construction, this particle is always put to the very end, and its scope covers all the clauses of the chain (25, 26; the limits of the scope are indicated by square brackets). Contrary to the other (TAM, negation) operators, the interrogative particle is localized outside the main clause, but this fact can hardly be regarded as invalidating the cosubordinating nature of the Bambara construction, cf. a similar situation in Barai language analyzed in (Foley & Van Valin Jr. 1984: 246).

- (25) *Ní màa-w nà-na só kà i ni à tó yèn,*
 if human-PL come-PFV.INTR house INF 2SG and 3SG remain there
[é ti í sigi sá kà à k̀nɔ wà?]
 2SG.EMPH IPFV.NEG REFL sit finally INF 3SG wait Q
 ‘If people come home and leave you there with him, won’t you sit down and wait for him?’ [Maningances].

- (26) [*Án ma bɔ́ án ka dùgu` lá kà sé é mà*
 1PL PFV.NEG exit 1PL POSS land\ART in INF arrive 2SG.EMPH ADR
yàn wà?
 here Q?
 ‘Haven’t we left our village and arrived at your place?’ [Contes bambara 1974].

The imperative in Bambara has a zero Aux for a singular subject, and for a plural subject the Aux is *yé*. As any other TAM marker, its scope covers the entire clause chaining construction.

- (27) *Á yé bɔ̀gɔ` nɔ̀ni, [á y` à tà k` à dí ní mà].*
 2PL IMP mud\ART dilute 2PL imp 3SG take INF 3SG give 1SG ADR
 ‘Dilute the mud, take it and give it to me’ [Cɛmandali Babilen].
- (28) *Á yé dúmuni` ké, kà mìnni` ké!*
 2PL IMP eating\ART do INF drinking\ART do
 ‘Eat and drink!’ [Kurane].

5.2. Scope of negation

Let us first overview the interaction of negation with constructions containing infinitives.

As mentioned in section 2, in Bambara, negation in a finite verbal clause is expressed by a predicative marker, jointly with TAM characteristics. In an infinitive construction, the predicative marker slot is occupied by the the infinitive marker *kà*; the only possibility to negate the action expressed by an infinitive is through negation of the matrix/head verb. An infinitive construction may be included or not included into the scope of negation of the matrix/head verb, depending on the type of syntactic relations between both verbal constructions.

It was already said in section 2 that if an infinitive depends on a modal matrix verb, the scope of negation includes both the matrix verb and the infinitive (29).

- (29) *Màa`-w ma sé kà kò` tìgɛ.*
 human\ART-PL PFV.NEG arrive INF river\ART cut
 ‘The people failed to cross the river’ [Jɛkabaara 171].

Dombrowsky-Hahn (2012: 55–56) interprets this fact in favor of the monoclausal character of the construction “modal verb + infinitive”, which seems reasonable. However, following Dumestre (2003), she does not distinguish between the groups of matrix verbs: “modal”, “aspectual” and “manner” verbs. As mentioned in section 2, in constructions of the aspectual verbs with infinitives, the scope of negation covers only the matrix verb, and not the infinitive (7b, 30). The impermeability of the matrix verb for negation seems to be a serious factor against the monoclausal character of this construction.⁸

⁸ On the problem of monoclausality vs. multiclausality of Bambara constructions, see 5.3.

- (30) *Ní í ma kɔ̀n k' ò tɔ̀n-na-don-wari in sàra,*
 if 2SG PFV.NEG precede INF that association-in-enter-money that pay
í tɔ̀gɔ̀ te sében.
 2SG name IPFV.NEG write
 ‘If you don’t pay this entrance fee in advance, you name won’t be included (in the list)’ [Kibaru 541] (i.e.: even if you pay the fee, but not in advance).

When an infinitive appears as a predicate of an adverbial clause of purpose, the scope of the negation expressed in the matrix clause includes both the matrix and the infinitive clauses. In reality, these constructions are similar to those with modal verbs, but they allow more easily pauses and adverbials after the matrix verbs, and there are therefore more reasons to regard them as biclausal.

- (31) *Á ma fɛ̀n sɔ̀rɔ̀ kà fɔ̀.*
 3SG PFV.NEG thing find INF say
 ‘He has found nothing to say’ [An ka yeɛ].
- (32) ... *à kó ní í ma fɛ̀ere ké kà ná ni filanin`-w yé*
 3SG say if 2SG PFV.NEG trick do INF come with twin\ART-PL PP
só` kɔ̀nɔ̀, à kó jí` te ná Dugufinnin.
 house inside 3SG say water\ART IPFV.NEG come Dugufinnin
 ‘He said: if you do not find a way to bring the twins to the house, water won’t come to Dugufinnin’ [Contes bambara 1974].

Table 1. Scope of negation in constructions with infinitives

Type of construction	Scope of negation
Modal verb + infinitive	entire construction
Aspectual verb + infinitive	matrix verb
Manner verb + infinitive	matrix verb or infinitive or both
Matrix verb + purpose infinitive	both clauses

Concerning the Bambara clause chaining constructions, one can find categorical judgments in special literature: “Such clauses cannot be negated; if they are, they are finite clauses and cannot be considered consecutive clauses anymore” (Dombrowsky-Hahn 2012: 56).

However, for my informants, negation is quite compatible with clause chaining; such examples are by no means exceptional in natural texts. A systematic examination of the occurrences of negative cosubordination⁹ shows that they can be divided into two groups:

⁹ I have made a search for the perfective negative marker *ma* followed by the infinitive marker *kà* (maximum 5 words to the right), then about 33% of set (about 320 examples, out of 964) have been sorted manually. In the obtained sample, about 8% represent clause chaining constructions.

- those (less numerous) with negation scope on the second verb (infinitive) (33, 34). In such cases we certainly have negation rising, which is cross-linguistically typical of subordinated clauses;
- those (much more frequent) where negation extends over the entire clause chain (35, 36).

The difference between these two types can be explained as follows. In the negation-rising clause chains, the actions of both clauses appear as simultaneous (ex. 33: God creates the fingers at the same time as he makes them equal; ex. 34: at the moment of Simbalan's death it grew clear that he left no children), while in the sentences of the second type (the negation scope covers both clauses) the actions are sequential. This explanation proves correct for the other examples of my sample as well.

When the infinitive clause is introduced by an alternative conjunction, the scope of negation necessarily extends on both clauses (37).

Among the natural examples of my sample, I have found no occurrences where negation scope would cover only the first clause.

- (33) *Ála ma bólokɔni`-w dá kà ù kénye.*
 God PFV.NEG finger\ART-PL create INF 3PL equalize
 'God has not made fingers equal' (= God has created fingers, he hasn't made them equal) [Sagesse bambara].
- (34) *Sinbalan ma sà kà dén tó.*
 Sinbalan PFV.NEG die INF child leave
 'Sinbalan died and left no children' (= he died, he has left no children) [Geste de Ségou].
- (35) *Háli bì, à ma kénye kà nà só fɔlɔ.*
 even today 3SG PFV.NEG recover INF come house yet
 'Until now, he has not yet recovered and hasn't returned home' [Chroniques amou-reuses].
- (36) *Màa sí ma kúma n fɛ, kà nà n jò Dikɔ*
 human nobody PFV.NEG speak 1SG by INF come 1SG stand Diko
ka biro` dá` lá...
 POSS office\ART mouth\ART in
 'Nobody spoke to me, nobody came to bring me to the door of Diko's office...' [Maningances].
- (37) *̀nka nê ma à t̩ɲetigiya wàlimà k' à nkàlontigiya.*
 but 1SG.EMPH PFV.NEG 3SG confirm or INF 3SG disprove
 'But I have neither confirmed nor disproved it' [Faba janjo].

So, in consecutive clause chains (which are the most current), negation behaves in the same way as the other operators tested above: its scope extends to the entire clause chain. Evidently, negation is attributed to the cosubordinate nexus, and each clause takes this specification there. Less frequent simultaneous clause chains behave, in this respect, like cross-linguistic subordinative sentences and require a negation rising.

5.3. The problem of mono- and pluriclausality

The problem of scope of negation has incidence on the highly controversial question of mono- and pluriclausality. It seems that the mainstream tendency today is to consider “the criterion of lack of independent negation as a cross-linguistically applicable test for clausehood” (Bohnenmeyer et al. 2007: 501). Haspelmath (2016: 299–301) argues strongly in favour of the «single negatability» as a universal criterion for monoclausality, and this approach grows popular among linguists.

According to this criterion, the consecutive constructions in Bambara should be regarded as monoclausal, in contradiction to other factors, such as the absence of a single intonative contour and possibility to have separate event-locational modifiers, as in (37).

In fact, Haspelmath assumes that the heuristic value of the term «monoclausality» is contestable:

Since clause status is not (yet) widely associated by linguists with single negatability, it would perhaps be more transparent if ‘monoclausal’ <...> were replaced by ‘singly negatable’. However, the term ‘monoclausal’ is less cumbersome, and it would be good if more linguists became aware that it actually has little meaning in a cross-linguistic context unless we apply the same definition in all languages (Haspelmath 2016: 301).

The Bambara data (in particular, the clause chaining constructions) is evidence for the separation of the single negatability from monoclausality. It is certainly just a question of labels: terms are conventional, and if specialists agree, they can use any term for any phenomenon. And still, it seems preferable to avoid conventions that bring us to too counterintuitive decisions, such as classification of Bambara clause chaining constructions with monoclausal ones.

6. Same- and different-subjectness

In Bambara clause chaining constructions, both same-subjectness and different-subjectness are allowed, but they are not distinguished by any morphological marker.

By default, all the clauses of a chain have the same subject, which is expressed only once, in the first clause (38).

- (38) *Dùgumɔ̀ɔ̀́ ye dúmuni dúman` tóbi, ø, kà súrukuba` fàna n' ò yé.*
 villager-ART PFV.TR food good\ART cook INF hyena\ART feed and it PP
 ‘The villagers cooked good food and fed the Hyena with it’ [Dunumba kumata].

Different-subjectness is much less frequent, but it is by no means exceptional (Dumestre 2003: 386). In such constructions, types of correspondence of the zero subject of the infinitive may vary.

Quite often, the verb of the full-fledged clause is in a passive construction,¹⁰ and the zero subject of the infinitive corresponds to the demoted (or eliminated) ex-subject of the main verb (39).

- (39) *Ní dénmisen-w_i kò-ra ø_? kà fini` dòn ù_i lá...*
 when child-PL wash-PFV.INTR INF clothes enter 3PL at
 ‘When children are washed and clothed...’ [Fasokan 2011_9].

- (40) *Né bólokò-la ø_? kà nê tó sà̀n tân ní fila lá.*
 1SG.EMPH circumcise-PRV.INTR INF 1SG.EMPH leave year ten and two at
 ‘I was circumcized when I was 12 years old’ (lit.: ‘I was circumcized, to leave me in 12 years’) [entretien sida 1994.04.09].

In other cases, no participant (even non-mentioned!) of the situation designated by the head verb can be regarded as the referent of the zero subject. So, in (41), the subject of the head verb is *Màli sòrodasi sàba ní`* ‘souls of three soldiers of Mali’. It is certainly not the referent of the zero subject of the infinitive. It is evident that the zero subject refers to the jihadists, but they are not mentioned overtly in the sentence. In (42), the situation is even more complicated: normally, verb *sà`* ‘die’ in Bambara is intransitive, but in the infinitive clause it appears as transitive, and the NP *à bá`* ‘his mother’ that should be, logically, its subject, appears as its direct object.

- (41) *Màli sòrodasi sàba ní` tó-ra à lá,*
 Mali soldier three soul-ART remain-PFV.INTR 3SG at
 ø_? *kà náani jógìn.*
 INF four wound

‘Three Malian soldiers were killed there and four were wounded’ (litt.: ‘Souls of 3 Malian soldiers remained in it...’) [Kibaru 496].

- (42) *Músà fã` sà-ra ø_? kà à bá` sà.*
 Musa father-ART die-PFV.INTR INF 3SG mother-ART die
 ‘Musa’s father died, and his mother died too’.

In synchrony, a “pseudo-causative consecutive infinitive” interpretation of this phenomenon advanced by Creissels (2009: 211) can be accepted: intransitive constructions should

¹⁰ As already said in section 2.1, in Bambara, the passive lability is absolutely regular: any formally transitive verb can be used in an intransitive construction, its ex-direct object being promoted to the syntactic function of subject, and its ex-subject being demoted to the status of an agentive complement (with a postposition *fè*) or eliminated from the surface level.

be postulated as initial (*{mògɔ}* *náani jógin-na* ‘four people got wounded’, *à bá` sà-ra* ‘his mother died’). In an infinitive construction, there is no slot for a subject; however, their subjects, being not co-referent with the subjects of the head verbs, cannot be simply omitted. This controversy is solved through an operation of “pseudo-causativation”: an indefinite causer (unexpressed at the superficial level) is introduced, and the ex-subject is demoted to the position of the direct object.

It should be specified that “pseudo-causative” constructions appear not only in clause chaining; they can be also introduced by subordinating conjunctions (43).

- (43) *Nê be tó í bá` ka kàlifa` kónɔ sánni kà*
 1SG IPFV remain 2SG mother\ART POSS protection\ART in before INF
sàya` sé í mà.
 death\ART arrive 1SG ADR
 ‘I will stay under your mother’s protection till my death’ (litt.: ... before the death reaches me’) [Bamanan nzirin].

7. A diachronic interpretation

Infinitive markers of the *kà* type (tone may vary) exist in all Manding varieties and in the Mokole language group. However, *kà* is not used as clause chaining marker everywhere; in particular, this function of *kà* is not attested in Mandinka.¹¹ This fact can testify for its relatively recent emergence, most probably, from the infinitive of purpose (Creissels 2009: 209). However, we can also try to search for the origin of the infinitive marker *kà*.

7.1. Odjenne Dyula evidence

Evidence from Odjenne Dyula (Braconnier 1991; Braconnier 1992) gives some clues for the etymology of *kà* as a clause chaining marker (and, probably, as an infinitive marker). In this variety, the non-initial clause of a different-subject consecutive construction may have the following structure: *ká S yè* (O) V (44).¹² In this construction, the predicative marker is *yè*, and the element *ká* is interpreted by Braconnier as a conjunction.

- (44) *Mósò-í náà dèn’ gbísí-rà ká cé-’-í yè gbísí fànà.*
 woman-PL PFV child-ART beat-INF2 CONJ man-ART-PL SEQ beat too
 ‘Women beat the child, and the men beat the child too’.

¹¹ On the other hand, *kà* is used for clause-chaining in Kakabe (Vydrin 2017: 146), a language of the Mokole group of Western Mande. It seems probable that this use of the infinitive marker could be explained by the influence of Maninka, a Manding language widely spoken by the Kakabe people as a L2.

¹² In Odjenne Dyula, tones are regularly inverted with respect to Bambara (and other Central Manding varieties). Therefore, *ká* regularly corresponds to *kà* in Bambara, and *yè* corresponds to *yé*.

In the same-subject constructions, the subject is omitted, and the sequence *ká yè* merges into *káà*:

- (45) *Áí mà kè káà ná.*
 3PL PFV.NEG do CONJ.SEQ come
 ‘They didn’t come’.

There are also cases where *ká* appears without *yè*; nevertheless, Braconnier considers it as a conjunction (*complémenteur*) in all the contexts.

7.2. A specialized consecutive marker *diyé* in Sikasso Bambara

In the Sikasso (Southern Mali) dialect of Bambara,¹³ there is a specialized consecutive marker *diyé*, in competition with *kà*. It can appear both in same-subject constructions, in which case it is preceded by no subject NP (46), and in different-subject ones, when the subject is overtly expressed (47).

- (46) *Sònsannin` ka dònsoke` dó yé ò yé,*
 hare\ART POSS hunter\ART one PFV.TR that see
diyé kólokari` fàga, diyé tága ò gòlo dí sònsannin` mà.
 SEQ wild.cat\ART kill SEQ go that skin give hare\ART to
 ‘One of Hare’s hunters saw this, he killed a wild cat, he went and gave its skin to the Hare’ [Contes Bambara 1974].
- (47) *Ò nà-na ò ké kà né kò n diyé sùnɔgɔ.*
 that come\PFV.INTR that do INF 1SG.EMPH wash 1SG SEQ sleep
 ‘She did it, she washed me, and I fell asleep’ [Contes Bambara 1974].

Although the origin of the initial element *di-* of the consecutive marker is obscure, the etymological identity of its final element *-yé* with the predicative marker *yè* in Odjenne Dyula seems highly probable. The existence of these markers in more than one Manding varieties may testify for the presence of a consecutive marker **yé* in Proto-Manding, replaced by *KA* in the majority of Manding varieties.

7.3. Origin of *kà*

The Ojenne Dyula evidence presented by Braconnier proves convincingly that in this Manding variety, the element *ká* in the consecutive construction is a conjunction, in the

¹³ Manding variety spoken in Sikasso and around this city is traditionally referred to as Dyula. However, in modern Mali, it is more often associated with Bambara. In any case, it belongs to the Manding dialect continuum that comprises Malian Bambara and Ivorian Dyula, where it would be difficult to trace a limit between the languages. To my knowledge, there is no grammatical description of the Sikasso Manding variety; the data discussed here are taken from texts of popular tales.

According to Aby Sangaré (p.c.), a similar consecutive marker exists in Dyula of Côte d’Ivoire.

course of transformation into a consecutive predicative marker through fusion with the original consecutive marker *yè*.

It cannot be excluded that this process concerns not only Ojenne Dyula, in other words, it may be a general evolution path for the entire Manding group. A hypothesis can be advanced that the infinitive marker *KA* in Manding (in all the contexts where the infinitive can appear) may be a conjunction by origin (and in Ojenne Dyula, we observe the final stage of its transformation into predicative marker).

If we assume the hypothesis of the conjunctive origin of the infinitive marker in Bambara, the weird different-subjectness (discussed in section 6) finds its explanation. What is now a direct object in “pseudo-causative infinitive constructions” was originally a true subject; it was converted into an object after grammaticalization of the conjunction into the infinitive predicative marker.

In Mande languages, it is very typical of the markers of non-finite verbal forms to go back to adpositions. As far as the predominant type of adpositions is represented by postpositions, the markers of the non-finite forms are usually suffixes (for example, the “second infinitive” suffix *-la* in many Manding languages going back to the postposition *lá*). However, in most Mande languages, there are also (not numerous) prepositions, which can also fulfil the conjunction function (Vydrin 2019: 169–173). Reflexes of a preposition/conjunction that could be the origin of the infinitive marker in Bambara are attested in Southwestern Mande languages: *gà* in Looma, as a preposition, expresses the instrumental, comitative, equative meanings; as a conjunction, it can introduce completive and purpose clauses; *à* in Kpelle (both a preposition and a conjunction), *à* in Mende and Loko (preposition), *ngàa* in Bandi (preposition).

One can also mention postpositions with instrumental, sociative, equative (and some other) meanings in South Mande languages: Dan *ká*, Mano *ká*, Goo *ká ~ kǎ*, Tura *gǎ ~ ǎ*. In Kla-Dan this postposition has grammaticalized into an infinitive suffix *-ká* (Makeeva 2017: 633, 650), therefore, this grammaticalization path (instrumental/comitative/equative adposition > infinitive marker) proves to be not unusual for Mande. It would be tempting to surmise a material cognacy between these South Mande postpositions and the Southwestern Mande prepositions, however, such a swap (from postposition to preposition, or vice-versa) would be unprecedented in Mande family, and should be therefore regarded as improbable (unless some serious pro arguments are found).

8. Conclusions

The clause chaining construction in Bambara displays some typologically unusual characteristics which deserve special mentioning.

Let us consider the following quotation from Longacre (2007: 417).

While medial-final chaining has a dominating final verb of fully inflected structure as opposed to medial verbs of defective structure, initial-consecutive structures have a dominating initial verb of one structure followed by consecutive verbs which are of different structure. Medial-final chaining is found in OV languages where it patterns as a further feature of head-final structures, while initial-consecutive chaining is found in VO languages (VSO and SVO) where it patterns as a further feature of head-initial structures.

Firstly, Bambara (and other Manding languages) contradicts the rule (taken by Longacre for granted) of a strong correlation between the VO basic order and the “initial-consecutive” type of clause chaining: the basic word order all over Mande family is SOV, nevertheless, the first verb is dominating, and the subsequent verbs are attributed a defective (infinitive) structure.¹⁴

Secondly, it is assumed by Longacre that the clause chaining constructions with a dominating initial verb are consecutive, and this opinion is shared by other linguists. As it was shown in section 4 (and also 5.2), it is not quite exact for Bambara: true, the consecutive meaning is the most current for clause chaining constructions, but it is not the only option available. A full-fledged verb can be indexed not only for the perfective aspect (prototypical for the consecutive meaning), but also for any other aspect and mode, and the actions of the subsequent verbs may be simultaneous with the action of the initial verb. Therefore, Longacre’s term “initial-consecutive chaining” is not applicable to Bambara (or applicable with reservations).

Bambara clause chaining construction behaves predictably with respect to the scope of traditional operators, such as illocutionary force and imperative: both extend their scope over the entire construction. The situation with the negation is more complicated. First, contrary to the existing opinion, clause chaining construction is compatible with negation. Secondly, there is a differentiation with respect to the scope of negation depending on the semantics of the clause chaining construction: in consecutive clause chains, the scope of negation covers the entire construction, while in simultaneous ones, negation (although indexed in the initial clause) has a scope on the non-initial clause (which represents a case of negation rising, more common for subordinative sentences).

Yet another peculiarity of the Bambara clause chaining construction is their differentiation by the feature of same-/different-subjectness. The different-subject infinitive clauses have no morphological marking, but they are characterized by a peculiar syntax. This syntax can be probably explained diachronically by the origin of the infinitive marker in a subordinate conjunction.

Glosses

1, 2, 3	1st, 2 nd , 3rd person	INTR	intransitive
ADR	adresative postposition	IPFV	imperfective marker
ART	tonal article (low floating tone)	NEG	negative
CERT	assertive future marker	PFV	perfective
COND	conditional marker	PL	plural
CONJ	conjunction	POSS	possessive marker
DISTR	distributive conjunction	PP	multifunctional postposition (comitative, instrumental, equative, etc. meanings)
EMPH	emphatic pronoun	PROG	progressive marker
FUT	future marker	PROH	negative subjunctive
ID	identification copula	PST	restrospective operator
IMP	imperative	PTCP	participle
INF	infinitive marker		

¹⁴ It should be noted however that, when speaking of correlation between the medial-final chaining and the OV word order, Longacre meant languages with the clause-final position of the verb (i.e., adverbials preceding the verb), while in Mande there is a basic word order S Aux O V X, where adverbials (designated by X) follow the verb.

Q	particle of general question	SBJV	subjunctive marker
RECP	reciprocal pronoun	SEQ	sequentative marker
REFL	reflexive pronoun	SG	singular
REL	relative determiner/pronoun	TR	transitive
RES	resultative		

Abbreviations

Aux	– auxiliary, predicative marker	TAM	– tense, aspect, mood
O	– direct object	V	– verb, verbal predicate
S	– subject	X	– oblique

Sources

- [Afiriki ka hadamaden] – *Afriki ka hadamaden ni fasojamaw ka josariyasun. Charte Africaine des Droits de l'homme et des peuples*. Translated by Bubakar Jara. Niamey: CELHTO, 1985.
- [An ka ye] – Dunbiya, Siyaka; Sangare, Bakari. *An ka yele dɔɔnin*. Bamako: Jamana, 1996.
- [Avenement de Da] – Avenement de Da. In: Dumestre, Gérard. *Geste de Ségou*. Paris : Armand Colin, 1979, pp. 265-357.
- [Baara kalan ka nyesin] – *Baara kalan ka nyesin balikukalan telimanw ma*. (Bamako) Koorisene baarada, 1985.
- [Balimamusoden] – Tarawele, Daramani. *Balimamusoden*. Bamako: Kalan Diya. n.d.
- [Bamako sigicogoya] – *Bamako sigicogoya*. Text collected and transcribed by Charles Bailleul.
- [Bamanan nzirin] – Kone, Kassim. *Bamanan nzirin Bamananankan ni Angilekan na (Bamanan tales in Bambara and English)*. West Newbury, Massachusetts: Mother Tongue Editions/An fasokan, 1995.
- [Bolociw] – *Bolociw*. Bamako: EDIM, 1977.
- [Cemandali Babilen] – Kone, Alikawu. *Cemandali Babilen*. Bamako: Jamana, 1989.
- [Chroniques amoureuses] – Dumestre, Gérard & Touré, Seydou. *Chroniques amoureuses*. Ms.
- [Cikela ni bara] – Fane, Baba. *Cikèla ni bara*. In : Annik Thoyer (éd.). *Le riche et le pauvre et autres contes bamanan du Mali*. Paris : L'Harmattan, 1997, pp. 123-132.
- [Contes bambara 1974] – Görög-Karady, Veronika & Meyer, Gérard. *Contes bambara du Mali et du Sénégal Oriental*. Paris, 1974
- [Contes bambara 1979] – Görög, Veronika. *Contes bambara du Mali*. Paris : Publications orientalistes de France, 1979.
- [Dɔnkesunguru] – Dɔnkesungurun. http://uz-translations.net/?category=afraudvid-african&altname=bambara_folk_tales_songs
- [Dunumba kumata] – Jara, Umaru Nanankoro. *Dunumba kumata*. Paris: Donniyakadi, 2011.
- [entretien sida1994.04.09] – *Entretiens sur le sida*. Zacharia avec ZD., 09.04.94 (transcript of an audio recording by Gérard Dumestre).
- [Faba janjo] – Berete, Hamidu. *Faba janjo. Maana ncinin nɔgɔndan sebenni bamananankan na (1993 nɔgɔndan)*. Bamako: Jamana, 1994.
- [Geste de Ségou] – Dumestre, Gérard. *Geste de Ségou*. Paris : Armand Colin, 1979.
- [Jekabaara] – *Jekabaara* (a monthly newspaper published in Bamako in 1986-2018).
- [Kibaru] – *Kibaru* (a monthly newspaper published since 1972 in Bamako).
- [Kolonkise 10] – *Kolonkise* (a periodical published in Bamako in 1990s).
- [Kurane] – *Kurane kɔɔ, bayelemanen bamananankan na*, Alihaji Modibo Jara bolo. (Bamako.), Vol. 1, 2011.
- [Manigances] – Dumestre, Gérard; Touré, Seydou. *Maléfices et manigances: Chroniques maliennes*. Ms.
- [Sagesse bambara] – Bailleul Ch. *Sagesse bambara : proverbes et sentences*. Bamako : Editions Donniya, 2005.
- [Sunjata ka maana] – Sidibe, Modibo. *Sunjata ka maana*. Transl. by Jelimakan Jabate. Bamako: Donniya, 2007.

References

- Bickel, Balthasar. 2010. "Capturing particulars and universals in clause linkage: A multivariate analysis". In: Isabelle Bril (ed.). *Clause Linking and Clause Hierarchy Syntax and Pragmatics*. John Benjamins, 51–101.
- Bohnenmeyer, Jürgen, Nicholas J. Enfield, James Essegbey, Iraide Ibarretxe-Antuñano, Sotaro Kita, Friederike Lüpke & Felix K. Ameka. 2007. "Principles of event segmentation in language: the case of motion events". *Language* 83(3), 495–532.
- Braconnier, Cassian. 1991. "Elements de syntaxe du verbe en dioula d'Odienne". *Mandenkan* 21, 1–140.
- Braconnier, Cassian. 1992. "Dernières nouvelles d'Odienné: «ka» infinitival mandingue est un «C», pas un «I» !" *Mandenkan* 23, 67–81.
- Creissels, Denis. 2006. *Syntaxe générale: Une introduction typologique*. 2 vols. Lavoisier: Hermes Sciences.
- Creissels, Denis. 2009. *Le malinké de Kita* (Mande languages and linguistics 9). Köln: Rüdiger Köppe Verlag.
- Dombrowsky-Hahn, Klaudia. 2012. "Motion events in Bambara (Mande)". *Journal of African Languages and Linguistics* 33(1), 37–66.
- Dumestre, Gérard. 2003. *Grammaire fondamentale du bambara*. Paris: Karthala.
- Foley, William A. 2010. "Clause linkage and Nexus in Papuan languages". In: Isabelle Bril (ed.). *Clause Linking and Clause Hierarchy Syntax and Pragmatics*. Amsterdam/Philadelphia: John Benjamins, 27–50.
- Foley, William A. & Robert D. Van Valin Jr. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.
- Haspelmath, Martin. 1995. "The converb as a cross-linguistically valid category". In: Martin Haspelmath & Ekkehard König (eds.). *Converbs in cross-linguistic perspective. Structure and meaning of adverbial verb forms – adverbial participles, gerunds* [Empirical Approaches to Language Typology 13]. Berlin: Mouton de Gruyter, 1–55.
- Haspelmath, Martin. 2016. "The serial verb construction: Comparative concept and cross-linguistic generalizations". *Language and Linguistics* 17(3), 291–319.
- Longacre, Robert E. 2007. "Sentences as combinations of clauses". In: Timothy Shopen (ed.). *Language Typology and Syntactic Description*, Vol. 2. Cambridge University Press, 371–420.
- Makeeva, Nadezhda. 2017. "Kla-dan jazyk (Кла-дан язык) [Kla-Dan]". In: Valentin Vydrin, Yulia Mazurova, Andrej Kibrik & Elena Markus (eds.) *Jazyki mira: Jazyki mande (Языки мира: Языки манде)* [Languages of the world: Mande languages]. St. Petersburg: Nestor-Historia, 617–679.
- Olson, Michael Leon. 1981. *Barai clause junctures: toward a functional theory of interclausal relations*. Australian National University, Ph.D. dissertation.
- Vydrin, Valentin. 2014. "Infinitiv v jazyke bamana (Инфинитив в языке бамана) [Infinitive in Bamana]". In: Mikhail Daniel, Ekaterina Ljutikova, Vladimir Plungian, Sergej Tatevosov & Olga Fedorova (eds.). *Jazyk. Konstanty. Peremennye. Pam'ati Aleksandra Jevgenjevicha Kibrika (Язык. Константы. Переменные. Памяти Александра Евгеньевича Кибрика)* [Language. Constants. Variables. In memoria Alexandre Kibrik]. St. Petersburg: Aleteia, 653–676.
- Vydrin, Valentin. 2016. "Tonal inflection in Mande languages: The cases of Bamana and Dan-Gweetaa". In: Enrique L. Palancar & Jean Léo Léonard (eds.). *Tone and Inflection: New facts and new perspectives* (Trends in Linguistics Studies and Monographs 296). Berlin: De Gruyter Mouton, 83–105.
- Vydrin, Valentin. 2019. *Cours de grammaire bambara*. Paris: Presses de l'INALCO.
- Vydrin, Valentin, Kirill Maslinsky & Jean-Jacques Méric. 2011–2020. *Corpus Bambara de Référence*. <http://cormand.huma-num.fr/index.html>.
- Vydrina, Alexandra. 2017. *A corpus-based description of Kakabe, a Western Mande language: prosody in grammar*. Paris: INALCO, Ph.D. dissertation.

Categorization of phasal polarity items in Bambara (Mande)

Abstract

The phasal polarity (PhP) operators ALREADY, NOT YET, STILL and NO LONGER may have different representation in particular languages. In the Mande language Bambara several items may be distinguished for each PhP expression. This paper discusses their categorization. The review of the most important dictionaries and grammars shows disagreement concerning their categorial status, which I attribute to different criteria on which the categorizations are based. Occurring in clause-final position, most items are assumed to be adverbs or particles. This paper provides a revised categorization of the PhP expressions using van Baar's (1997) criteria, especially the principles he proposes to distinguish adverbs and particles. It is assumed that the PhP items in Bambara attest different stages of grammaticalization. Some PhP items are adverbs, others are grammaticalized particles; however, showing little formal reduction, they occur in the same form as the elements supposed to be their sources.

Keywords: phasal polarity, adverbs, particles, grammaticalization

1. Introduction

The morphosyntactic analysis of texts and individual sentences requires the analyst to gloss individual items, a process often associated with the determination of the word class they belong to¹. This task is usually straightforward when the items to gloss are full lexical words; however, it is more subtle when it comes to items situated somewhere on the continuum between lexical and grammatical elements. Phasal polarity or, henceforth, PhP expressions, which include the notions of ALREADY, NOT YET, STILL and NO LONGER come

¹ After a presentation on PhP expressions in Bambara that I gave at the International Conference on *Phasal Polarity in Sub-Saharan African Languages* at the University of Hamburg in February 2018, a colleague suggested I should gloss the PhP expressions using the source concepts, for instance *bán* as 'finish' instead of 'not yet'. This remark gave the incentive to study the categorial status of the PhP expressions. The outcome was presented at the 3rd SyWAL, which took place from 26 to 29 September 2018 in Warsaw. I would like to express my sincere gratitude to the organizers and participants of the symposium for their stimulating questions and remarks, and to the editors of this volume for their suggestions, comments and remarks that helped to improve the quality of this paper. Of course, all remaining shortcuts are my responsibility. The journey to Warsaw was made possible thanks to a grant supporting women in science and research (*Stabsabteilung Chancengleichheit*) at the University of Bayreuth. I am indebted to my language assistants Aisé Touré and M. Traoré for their help and patience.

in a huge variety of forms across languages: adverbs, particles, verbs within a construction consisting of several verbal components, verbal affixes used for the expression of tense, mood and aspect, or auxiliaries (cf. van Baar 1997: 41). For van Baar (1997: 214) such a variety of words belonging to different word classes and morphemes suggests that PhP expressions have different sources and attest different stages of grammaticalization, the process whereby lexical items of theoretically every word class can develop into grammatical elements.

This paper examines the categorial status of PhP expressions in the Mande language Bambara spoken in Mali. PhP is overtly expressed, i.e. Bambara has a full inventory of phasal polarity items. They occur in most cases in the clause final position, the preferred place for adverbs and particles. The study of Bambara PhP expressions in published dictionaries and descriptions reveals disagreement concerning their categorial status. The authors are either not unanimous in categorizing them as adverbs, particles or “elements similar to particles”, or they do not categorize them at all. I re-examine the categorial status of the Bambara PhP expressions using the criteria for the distinction between adverbs and particles suggested by van Baar (1997).

The article is structured as follows: in section 2 I introduce the definition of phasal polarity, which is followed by a short summary of morphosyntactic facts of Bambara (section 3); in section 4 I provide the inventory of PhP items in Bambara; the categorial status of each of them is discussed in section 5. A summary and concluding remarks are found in section 6.

2. Definition

Phasal polarity (PhP) is the linguistic realization of successive states of being and non-being, and phasal polarity items are “structured means of expressing polarity in a sequential perspective” (van Baar 1997: 40). They include *ALREADY* and *STILL*-expressions occurring in affirmative clauses, and *NOT YET* and *NO LONGER*-expressions, occurring in negative clauses. A phasal polarity item is used when the speaker connects a situation or the absence thereof to a previously or a posteriorly occurring opposite situation. The opposite situation is presupposed: it has either been mentioned in the discourse, or it exists in the speaker’s (and the addressee’s) imagination. The *ALREADY*-expression states that the situation holds at reference time implying a reference point at a previous moment where the situation doesn’t hold. The *STILL*-expression shows that a state holds at reference time relating it to a subsequent period where it doesn’t hold. *NOT YET* indicates that a situation does not hold at reference time relating it to a subsequent moment at which it holds. The *NO LONGER*-expression states that a situation does not hold referring it to a previous moment at which it holds.

Due to their semantics, PhP expressions “can be distinguished from other expression types in the language” (van Baar 1997: 41) and therefore constitute a structured set of forms. In his typological study, van Baar (1997: 214) makes a twofold observation concerning the form and the categorization of PhP expressions. First, different languages use various forms for the expression of one and the same notional category. For instance, in Hausa the notion of *ALREADY* is manifested as a verb-like expression showing all characteristics of the verb *rigā* ‘to precede’; in Burmese the same notion is represented by a clause-final

verbal suffix used for the expression of tense, mood and aspect; and in West Greenlandic, ALREADY is an aspect marker affixed to the verb with a position in the verbal string varying according to its scope. Moreover, the PhP expressions existing in an individual language are not necessarily attributed to one word class (part-of-speech in van Baar's terms). The author lists the part-of-speech membership of PhP expressions in some of his sample languages (van Baar 1997: 363). For instance Ewe ALREADY *xóxó* is an adverb, STILL *ga* is a particle, and the membership of NOT YET *hadé* could not be determined by van Baar (1997: 363, 74, 76, 81).

Bambara is a language in which the assignment of the PhP expressions to word classes is troublesome, as the existing categorizations show. According to Dumestre, *kàban* as ALREADY 1 expression, i.e. 'already' used in neutral scenarios, is a petrified construction based on a verbo-nominal lexeme (Dumestre (2003; 2011a; 2011b), an adverbial phrase according to Bailleul (2007: 196), and an adverb according to *Bamadaba* and Kone (2010). Second, there are items (*túguni* 'still, no longer', *bán* 'not yet') that have been assigned to different categories in different publications by one and the same author (first as an adverb, cf. Dumestre 1987: 493, later as an element close to particles, cf. Dumestre 2011a: 1003, 83). Third, other authors omit mentioning the grammatical category of certain items (i.e. *bílen* 'still, no longer' and *bán* 'not yet' (Bailleul 2007)) altogether.

Before providing the PhP system in Bambara and proposing an alternative categorization of its items, I discuss the language's morphosyntax indispensable for the understanding of the main section of this paper.

3. Notes on Bambara morphosyntax

Bambara shows S(O)V(X) word order. X refers to obliques and more peripheral elements. Auxiliaries that follow the subject NP, generally called predicate markers in Mande language studies, are obligatory constituents of a clause; they are constitutive of the TMA-system of the language. The TMA-system distinguishes basically a perfective, an imperfective and a progressive aspect and two future tenses. Predicate markers or verbal suffixes are portmanteau morphemes that cumulate a TMA value and polarity. Table 1 resumes the affirmative predicate markers and the single existing suffix, and the corresponding negative morphemes appearing in clauses with a verbal predicate.

TAM	affirm.	gloss	neg.	gloss
transitive perfective	<i>ye</i>	PFV.TR.AFF	}	PFV.NEG
intransitive perfective	<i>-ra--la--na</i>	PFV.ITR.AFF		
imperfective	<i>be</i>	IPFV.AFF	<i>te</i>	IPFV.NEG
future (intentional ²)	<i>bénà</i>	FUT.AFF	}	FUT.NEG
future (assertive)	<i>na</i>	FUT.ASS.AFF		
subjunctive	<i>ka</i>	SBJV.AFF	<i>kàná</i>	SBJV.NEG

Table 1. Affirmative and negative predicate markers in clauses with a verbal predicate

² Idiatov (2000) calls this value also "deontic or prospective."

In transitive clauses, predicate markers are inserted between the subject and the object NPs (1), (2). In intransitive clauses, all TMA values but the perfective affirmative are marked by predicate markers that follow the unique obligatory nominal term (3); the perfective affirmative is replaced by a suffix *-ra* (or one of its allomorphs *-la*, *-na*) (4).

- (1) *cè-* ` *ye* *mùso-* ` *wéle*.
 man-ART PFV.TR.AFF woman-ART call
 ‘The man has called the woman.’
- (2) *mùso-* ` *ye* *cè-* ` *wéle*.
 woman-ART PFV.TR.AFF man-ART call
 ‘The woman has called the man.’
- (3) *dén-* ` *kàna* *kàsi*.
 child-ART SBJV.NEG cry
 ‘The child shouldn’t cry ~ May the child not cry.’
- (4) *dén-* ` *kàsi-ra* (*só-* ` *kónɔ*).
 child-ART cry- PFV.ITR.AFF house-ART PP
 ‘The child cried (in the house).’
- (5) *à* *màga-ra* *ń* *bólo-* ` *lá*.
 3SG touch-IPFV.AFF 1SG arm-ART PP
 ‘It touched my arm (lit. touched at my arm).’

As demonstrated in (1), (2), grammatical relations are indicated by word order alone; there is no case marking, nor is there any indexation of the subject or the object on the verb. Postpositions mark oblique arguments (5) and more peripheral terms indicated in parentheses in (4); there are also some prepositions likely to introduce more peripheral terms.

Next to clauses with verbal predicates, Bambara has several types of clauses with nonverbal predicates. Such predicates are nouns or NPs, adjectives (Dumestre (2003: 171), called “quality verbs” by Vydrin (2019: 69) and “state verbs” by Creissels (1985)), and locative phrases. In these clauses, distinct copulae connect the subject to the predicate, as illustrated in (6)-(9). The presentative clauses with *dòn* (neg. *té*) lack the bipartition into subject and predicate (6).

- (6) Presentative clause
 a. *fàli-* ` *dòn*. b. *fàli* *té*.
 donkey-ART COPID donkey COP.NEG
 ‘This **is** a donkey.’ ‘This **is not** a donkey.’
- (7) Equational clause
 a. *ń* *dénke-* ` *yé* *kàlanden* *yé*.
 1SG son-ART COPEQU student PP
 ‘My son is a student.’

- b. *ń dénkɛ-` té kàlanden yé.*
 1SG son-ART COP.NEG student PP
 ‘My son is not a student.’

(8) Locative / existential clause

- a. *wáraba-` bé kúngo-` lá.*
 lion-ART COPLOC wilderness- ART PP
 ‘The lion is in the wilderness. ~ There is a lion in the wilderness.’

- b. *wáraba-` té kúngo-` lá.*
 lion-ART COP.NEG wilderness- ART PP
 ‘The lion is not in the wilderness.’

(9) Qualifying clause

- a. *fòronto-` ká fárin.* b. *fòronto-` mán fárin.*
 chili-ART QUAL.AFF hot chili-ART QUAL.NEG hot
 ‘The chili **is** hot.’ ‘The chili **is not** hot.’

The copulae are listed in the following table together with their negative counterparts:

clause type	affirm.	gloss	neg	gloss
presentative	<i>dòn</i>	COPID		
equational	<i>yé</i>	COPEQU	} <i>té</i>	COP.NEG
locative	<i>bé</i>	COPLOC		
qualifying	<i>ká</i>	QUAL.AFF	} <i>mán</i>	QUAL.NEG

Table 2. Copulae in clauses with nonverbal predicates

The locative type of clause is the source concept of several other, more abstract concepts, such as predicative possession (10) and physical (11) or mental experience constructions (12).

- (10) *wári-` bé ń bólo.*
 money-ART COPLOC 1SG PP
 ‘I have money.’

- (11) *kóngo-` b’ ù lá.*
 hunger-ART COPLOC 3PL PP
 ‘They are hungry.’

- (12) *à mĩiri-` té sáraka-` lá bilen.*
 3SG thought-ART COP.NEG sacrifice-ART PP no.longer
 ‘She does no longer think about the sacrifice’

Clauses with nonverbal predicates have present time reference, unless they are marked by the inactuality operator *tùn* or, especially in narrative texts, when an introductory sentence is marked by *tùn* locating the entire situation into a context at past time. *Tùn* can also appear in verbal clauses. In independent clauses, it produces past meaning; in the subordinated clauses of complex sentences, it creates a pluperfect reading.

The presentative/existential or the locative copula is used together with the perfective participle³ *V-len* to form resultative or perfect aspect (13).

- (13) *fɔlɔ-` jà-len b'é, jí fóyi té à lá*
 first-ART dry-PTCP COPLOC 1SG not.any COP.NEG 3SG PP
 [The child saw three wells]. The first was dry, there was no water at all in it.
 (04dinye_yaalala.dis.html)⁴

Further aspectual values such as the progressive, continuative, iterative, prospective, and modal values such as the deontic and obligative are obtained by means of periphrases based on the basic clause types listed above or by the use of auxiliaries grammaticalized to variable degrees (cf. Dumestre 1999; Idiatov 2000: 39-44).

The infinitive, which corresponds to the citation form of a verb bears the infinitive morpheme *kà* (*k'* before a pronoun with initial vowel). In the citation form of intransitive verbs, the verb immediately follows the infinitive morpheme (14a); transitive verbs are preceded by a pronoun replacing a possible object argument (14b), and reflexive verbs by the pronoun of the second person singular standing for a pronoun coreferential with the subject referent (14c).

- (14) a. *kà síran* b. *kà à bùgɔ [k'à: bùgɔ]* c. *kà í sīgi*
 'be afraid' 'hit him/her/it' 'sit down'

The infinitive takes part in a series of constructions, among others, consecutive constructions. In the latter, only the first clause is complete, bearing a subject and a predicate marker or verbal suffix, while the consecutive clause is introduced by the infinitive morpheme (15).

- (15) *í b' ò bɔ k' ò fili*
 2SG IPFV.AFF AN take.out INF AN throw.away
 'You take it outside and throw it away'

³ "Resultative participle" in Vydrin's terms (Vydrin 2019).

⁴ All source indications ending in .html are taken from the Bambara Reference Corpus (Vydrin et al. 2011-2018).

4. PhP-expressions in Bambara

4.1. The inventory

Bambara PhP expressions are independent clause-final words illustrated as follows.

- (16) *jàgo*-*minanw* *nà*-*na* *kàban* *wà*?
 trade-goods come-PFV.ITR.AFF already Q
 ‘Have the goods already arrived?’ (dukure-ni_san_cyenna.dis.html, #409098)
- (17) *ù* *má* *nà* *fólo*.
 3SG PFV.NEG come not.yet
 ‘They have not come yet.’
- (18) *ò* *sìgikafo* *bé* *sèn* *ná* *háli* *bì*.
 DEM dialogue COP.LOC foot PP even today
 ‘This dialogue is still going on.’
 (jekabaara271_05sidibe-mali_kenyekafekeleko-zup.repl.html, #2246840)
- (19) *ù* *té* *síran* *à* *né* *tún*
 3PL IPFV.NEG be.afraid 3SG PP no.longer
 ‘They are not afraid of it anymore.’ (bamanankan_kalangafe6-zup.repl.html, #295353)

A fuller set of PhP expressions is shown in figure 1. The numbers 1 and 2 stand for expressions used in pragmatically distinct contexts: ALREADY 1, NOT YET 1 occur in neutral contexts, whereas ALREADY 2, NOT YET 2, etc. appear in counterfactual contexts, i.e. in a situation with a certain counterexpectation (cf. Dombrowsky-Hahn to appear). We find two items for the neutral scenarios of ALREADY, NO LONGER respectively, and even more items for the neutral STILL and NOT YET scenarios. Only the ALREADY-expressions differ in neutral (*kàban*) and counterfactual scenarios (*kélèn*), whereas the more marked NOT YET and STILL-expressions are encoded by the cumulation of the PhP-element used in the respective neutral scenario and another item of the PhP-system.

ALREADY 1	<i>kàban</i> (<i>kàkòrò</i>)	NO LONGER 1	<i>bilen</i> <i>túgun</i>
ALREADY 2	<i>kélèn</i>	NO LONGER 2	
NOT YET 1	<i>bán</i> <i>fólo</i> <i>háli bì</i>	STILL 1	<i>háli bì</i> <i>háliisa</i> <i>túgun</i> (<i>bilen</i>)
NOT YET 2 (,still not‘)	<i>háli bì ... bán</i> <i>háli bì ... fólo</i>	STILL 2	<i>háli bì ...bilen</i>

Figure 1: PhP-expressions in Bambara

All simple (not cumulated) items occur clause-finally, either immediately following the verb or after a postpositional oblique or an adjunct if such an element comes after a verb, nevertheless preceding the interrogative particle *wà*, which marks yes/no questions (16). Exceptions are the STILL expressions *háli bi* and *hálisha*, which can figure either in clause final or clause-initial position. They obligatorily occur in clause-initial position when they are part of the more marked NOT YET 2 and STILL 2 expressions. The other element (*bán*, *fɔ́lɔ* or *bilen*) terminates the sentence in this case.

Since the word order is absolutely strict in Bambara, there are only very few word classes that can fill the indicated final position, chiefly adverbs and particles, suggesting that the PhP expressions represent one of these two word classes. However, due to different criteria underlying the categorization of items to word classes, the most important Bambara dictionaries and grammars disagree about the question which word class the PhP items belong to.

4.2. Criteria for the categorization of PhP items in earlier works

To classify particles, Dumestre (1987, 2003, 2011a, 2011b) uses the criteria of scope, function, distribution, and compatibility with nominal use. For Dumestre (2003: 303), a particle is a grammatical element that has either a single term (or nominal constituent (2011b: 4)) or the entire sentence in its scope. A particle can be omitted without invalidating the sentence. The main function of a particle is to insert the message content into the discourse. Dumestre (2011b) divides particles into two groups. The first one (“*particules contrastives*”) groups the so called contrastive particles, defined as elements that have scope over a nominal or an entire “segment” (clause, sentence). With the second group he associates sentence particles, i.e. elements that have scope only over an entire sentence (“*particules phrastiques*”). Particles and adverbs are both optional, but they differ in their scope: adverbs are basically related to the predicate, while sentence particles have scope over the entire sentence.

Dumestre very generally postulates for many Bambara items membership in two or more word classes such as the membership in the class of verbs and in the class of nouns, which he calls *verbo-nominals* (i.e. *bàna* ‘sickness’, ‘be sick’), or the membership in the class of determiners, in the class ‘close to sentence particles’ and *verbo-nominals* (*fɔ́lɔ* ‘first’, ‘not yet’, ‘do first’, ‘start’, ‘former times’).

In his most recent publication, the dictionary, Dumestre (2011a) classifies the PhP expressions as follows: only *hálisha* is a (complex) adverb, *bilen* ‘still’ and ‘no longer’ are contrastive particles, *túguni* ‘still’ and ‘no longer’, *bán* ‘not yet’ and *fɔ́lɔ* ‘not yet’ are close to sentence particles, *kélɛ́n* ‘already’ is a noun, *kában* ‘already’ and *kàkɔ́ɔ* ‘already’ figure under the entry of the *verbo-nominals* *bán* ‘finish’ and *kɔ́ɔ* ‘be, get old’ respectively; *háli bi* is considered to be a contrastive particle + a noun.

Vydrin (2019) uses mainly morphosyntactic criteria to determine the word class an item belongs to. Morphological principles refer to the potential of an item to combine with auxiliary words and the type of derivation it can be subject to. Syntactic criteria refer to the distribution of the items, i.e. their capacity to fill particular slots in a sentence. Giving preference to these two sorts of formal criteria, he takes into consideration semantic criteria only if there is a distinction between primary and secondary functions of a lexeme (Vydrin

2019: 301). Thus, by applying distributional criteria, the author considers the element *kélèn* ‘already’ – identified here as a PhP expression – as postverbal adverb, which differs from the numeral *kélen* ‘one’ only tonally. According to Vydrin, *kàbán* ‘already’ and *bán* ‘not yet’ (French: ‘encore’, which is in fact translatable as ‘still’ or ‘not yet’), two other PhP expressions, are postverbal adverbs too; they are derived by conversion from the verb including or excluding the infinitive morpheme *kà*, *kà bán* ‘to finish’ or *bán* ‘finish’ (Vydrin 2019: 371, 305). All three figure among the “spatial and temporal adverbs” (Vydrin 2019: 371).

The author (Vydrin 2019: 521) considers particles as elements that are not constituents of sentences; they have rather additional roles, especially modal functions. Particles differ from adverbs in not participating in the construction of denotational meaning of noun phrases or sentences. They differ from interjections by their inability to form sentences, i.e. they are obligatorily connected to a syntactic constituent such as a word, a phrase or a sentence. Three of the items recognized here as PhP expressions are categorized as particles by Vydrin (2019: 526, 529, 528): *bilen*, *háli bi* and *hálisha*.

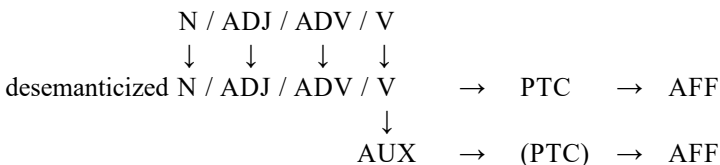
Bailleul (2007) and Kone (2010) do not reveal the criteria used for the assignment of the dictionary entries to specific word classes, and it is in these authors’ dictionaries that several PhP expressions figure without indication of word classes they belong to. This procedure is not restricted to PhP expressions, though. Bailleul classifies *háli*, which occurs in the PhP expressions *háli bi* and *hálisha* as a particle; *túguni* ‘still’ and ‘no longer’, and *fóló* ‘not yet’ are assigned to adverbs, *kàban* ‘already’ is considered to be an adverbial phrase. For Kone (2010), *kàban* and *kélèn* ‘already’ are adverbs, *kàkərɔ* ‘already’ is a compound noun, *bilen* ‘not yet’ and ‘still’ are adverbs; the author does not assign the other items to a category⁵.

In the following I present van Baar’s thoughts on particles and his criteria for the distinction between particles and adverbs, which I use subsequently for a re-categorization of the Bambara PhP expressions.

5. A new approach to the categorization of Bambara PhP expressions

5.1. Van Baar’s criteria for the distinction between particles and adverbs

According to van Baar (1996; 1997: 244, 230), particles are highly grammaticalized words and/or clitics that have lost the essential characteristics of their source categories. Diverse major word classes are possible sources of particles – nouns, adjectives, adverbs, or verbs, schematized as follows (Baar 1997: 229).



⁵ In Kone’s (2010) monolingual dictionary, the word classes are named as follows: *sementiyalan* ‘adverb’; *təgɔdorogolen* ‘compound noun’. Kone considers *halɪ* as a *nɛbila* ‘preposition’, but he does not indicate the word classes for *halibi* and *halisa*, which he writes as one word respectively.

Particles present particular features at different levels. First, they show reduction on the phonological level: polysyllables are reduced to monosyllables; clusters or long phonemes are simplified to simple sounds; segmental phonemes are brought down to suprasegmental phonemes; tonal tunes are simplified and the ability to mark distinctive stress gets lost (van Baar (1997: 231). Losing phonological weight, particles tend to evolve from independent words to clitics and, further, to affixes (AFF). As linkers rather than building blocks of referents, particles are unlikely to be the head of a construction or to be inflected (van Baar 1996: 266). Particles “only become meaningful if they are used in a specific context, i.e. they do not have a conceptual meaning, but rather a contextualized function” (van Baar 1997: 235). Since they are functional expressions, they are incapable of being involved in word formation processes which have lexical, i.e. context-independent meaning.

The main function of particles is “bridging between (part of) the sentence and a wider setting” (van Baar 1996: 290). The bridging takes place either between a textual unit and a wider textual setting, or between a textual unit and a nontextual setting. Considered to be a kind of relator, a particle is unlikely to be isolated from its relatum. Particles are decategorized; i.e. they do not show the properties of their sources. Particles derived from verbs do not have predicative use; those derived from nouns are not heads of a noun phrase; those derived from an adjective cannot be used as a modifier of a nominal head; those derived from an adverb cannot be used as a modifier of a non-nominal head. However, one item can belong to different categories depending on the context in which it is used and the formal characteristics related to this use (van Baar 1997: 242).

Van Baar (1997: 243) adopts the prototype approach⁶ for the identification of items as particles: the more features defining particles an item shows, the more prototypical representative of this class it is. An item that bears all features is a fully grammaticalized particle; an item that shows some but not all features, is a less ideal candidate of this word class. However, the criteria are not all weighted the same. Only three criteria seem to be really decisive: 1) the isolation criterion is the most important of all: only those items that cannot be isolated count as particles; 2) they must be unable to function as head of a construction; 3) they must be decategorized to a sufficient degree.

Van Baar (1996: 277ff; 1997: 244) distinguishes particles from adverbs, which are formally and distributionally similar to them. Both are often related to a sentence as a whole; further, both are capable of expressing PhP. However, he admits that the differences between the two categories are very subtle in a large number of cases. An adverb (or adverbial predicate in van Baar’s (1997: 241) terminology) is defined as an element that can be used as modifier of a non-nominal head, without further measures being taken. Thus, if an item is compatible with a noun, it is not an adverb. Particles, however can collocate with nominals, but without showing adjective-like characteristics, for instance number agreement.

Second, they differ with respect to their autonomy. Thus, particles are not isolatable from the rest of the sentence, whereas adverbs usually are. For instance, the English adverb can form a single word utterance (20a), whereas a particle cannot (20b).

- (20) a. *When will he come?* – *Tomorrow.*
 b. *(When) will he come?* – **Still*

⁶ Using this approach, van Baar follows Heine (1993) who applies it to auxiliaries.

Another peculiarity distinguishing particles from adverbs is the loss of semantic integrity, a term that van Baar adopts from Lehmann (1982) and that he defines as follows. “The function/meaning of particles is highly determined by the context, whereas the meaning of adverbials, for example, is autonomous to a high degree” (van Baar 1996: 280). The author names several factors that decide over the autonomy of items. First, the influence of negation on the item; second, the link between its position and its function. As a third factor, van Baar (1997: 248) names the relation of an item to the context.

Feature	Adverbs	Particles
compatibility / scope	used as modifier of a non-nominal head	particles and nominals can collocate, but no adjective-like characteristics
autonomy	autonomous: isolatable	not autonomous: not isolatable
semantic integrity	meaning is autonomous to a high degree, i.e. it remains the same under negation, in different position, in different context	loss of semantic integrity: function /meaning of particles is highly determined by the context (negation, position and wider context affect the function of a particle)

Table 3: The difference between adverbs and particles following van Baar’s criteria

5.2. Reconsideration of PhP-items in Bambara using van Baar’s criteria

The clause-final position of PhP items suggests that they are adverbs or particles, which are also the categories they were assigned to most frequently. In the following I will examine the Bambara PhP expressions using van Baar’s criteria to distinguish adverbs from particles, namely 1) the degree of their autonomy; 2) the degree of semantic integrity, i.e. the influence of negation and position or other contextual factors on their meaning or function. None of the items shows adjective-like characteristics such as plural marking, therefore this criterion does not need to be studied separately. Since van Baar (1997: 244, 230) considers particles to be strongly grammaticalized items, their possible sources and the degree of grammaticalization they have attained respectively is decisive for the determination of their membership in a word class. Therefore 3) their form and 4) their possible sources and the degree of their decategorization will also be discussed.

PhP-expression	PhP-item	Other meanings/functions of the items in Bambara
ALREADY1	<i>kàban</i>	<i>bán</i> ‘terminate’
	<i>kàkɔɔ</i>	<i>kɔɔ</i> ‘be/get old’; ‘old’; ‘elder sibling’
AREALDY2	<i>kélèn</i>	<i>kélen</i> ‘one’; ‘the same’, ‘alone’, ‘to the exclusion of’
STILL	<i>háli bì</i>	<i>háli</i> ‘even’, <i>bì</i> ‘today’
	<i>háli sa</i>	<i>háli</i> ‘even’, <i>sa</i> ‘now’
	<i>túgun</i>	<i>túgun</i> ‘again’, ‘no longer’
	<i>bìlen</i>	<i>bìlen</i> ‘again’, ‘no longer’
NOT YET	<i>fɔlɔ</i>	<i>fɔlɔ</i> ‘first’, ‘at present’, ‘do first’, ‘the first’

	<i>bán</i>	<i>bán</i> ‘terminate’
NO LONGER	<i>túgun</i>	<i>túgun</i> ‘again’, ‘still’
	<i>bilen</i>	<i>bilen</i> ‘again’, ‘still’

Table 4. Representation of the PhP expressions in Bambara and other meanings they denote or functions they have in the language.

5.2.1. *kàban* ‘already’

The source of the ALREADY item *kàban* is the verb *bán* ‘finish, terminate’. Its use as an independent verb is illustrated in (21). The PhP-expression developed from a more complex conceptual entity, the consecutive construction⁷, containing the verb *bán* in final position. It is introduced by the infinitive morpheme *kà*, which signals reference identity with the subject of the preceding full clause. *bán* ‘finish’ is likely to function as intransitive or as transitive verb. The transitive use of the verb in a consecutive construction requires the presence of a pronoun coreferential with the object of the preceding verb and of *bán* inserted between the infinitive morpheme and the verb (*à* in (22), coreferential with *sògo*, the object of the verb, *dún* ‘eat’). Usually, the vowel of the infinitive morpheme *kà* elides, leaving the object pronoun lengthened.

(21) Independent verb

à bán-na wà?
 3SG finish-PFV.ITR.AFF Q
 ‘Is it finished?’ (Dumestre 2011a: 83)

(22) Verb in consecutive construction

dùga-w yé sògo dún kà à bán [kà:bá].
 vulture-PL PFV.TR.AFF meat eat INF 3SG finish
 ‘The vultures ate the meat and finished it.’

(23) Particle in final position of a transitive clause

dùga-w yé sògo dún kàban [kàbá].
 vulture-PL PFV.TR.AFF meat eat already
 ‘The vultures have already eaten the meat.’

(24) Particle in final position of an intransitive clause

ń kòromuso sàba bée fúru-la kàban ...
 my elder.sister three all get.married-PFV.ITR.AFF already
 ‘All my three elder sisters are already married ...’ (Touré 1999: 76)

However, as PhP expression, the vowel is realized short, whether it is placed at the end of a transitive (23) or an intransitive (24) clause. This shortening can be interpreted as formal reduction constitutive of the grammaticalization process. Additionally, it is written

⁷ Like in the case of grammaticalization of other units, for instance auxiliaries, the process includes a concrete, schematic content employed for the expression of abstract grammatical concepts (Heine 1993:86).

as one word by Bambara speakers themselves (cf. Kone (2010: 90), Touré (1999: 76), cf. example 24), and also in the Bamadaba online dictionary, suggesting that it has become a word on its own.

There are some contexts showing that the item *kàban* has reached a rather high degree of grammaticalization as a PhP expression. First, it lost its verbal meaning, for it can be used with the same item as the main, introductory verb preceding the consecutive clause without yielding durative meaning (25); second, the PhP expression has been generalized to a certain degree, for it occurs in clauses with nonverbal predicates, which usually are not likely to be added consecutive constructions⁸. (26) illustrates a locative (hence nonverbal) predicate. However, the PhP item maintains the infinitive marking which is unequivocally a verbal feature.

- (25) *filimu bân-na kàban.*
 film finish-PFV.ITR.AFF already
 ‘The film is already finished’.

- (26) *móbili bé Musa fê kàban.*
 car COPLOC Muss PP already
 ‘Musa has a car (lit.: a car is with Musa) already’.

It is impossible to isolate *kàban* in its meaning of ‘already’ from the remaining part of the clause, providing evidence that it differs from regular adverbs like *kúnùn* ‘yesterday’.

- (27) *Q: à nà-na wágati jùmen? A: kúnùn.*
 3SG come- PFV.ITR.AFF time which yesterday
 ‘When did s/he come?’ ‘Yesterday.’

- (28) *Q: à nà-na wà? A: *ɔwɔ, kàban.*
 3SG come-PFV.ITR.AFF Q yes already
 Has s/he come? Yes, already.

Kàban is specialized as PhP expression ALREADY only in affirmative clauses. It did not further evolve as NOT YET expression, maintaining features of its source, the consecutive construction, which follows an affirmative finite clause in the majority of occurrences⁹.

- (29) **à má nà kàban.*
 3SG PFV.NEG come already
 S/he has not come already/yet.

⁸ There are, however, other highly grammaticalized expressions showing the infinitive *ká*, for instance *kà bɔ́*, which has developed into a kind of preposition meaning ‘from, out of’ (cf. Dombrowsky-Hahn 2012). Dumestre (2003: 400-401) further mentions other particular cases where an infinitival phrase can follow a clause with a nonverbal predicate.

⁹ Consecutive constructions following negative clauses are much less common.

There are contexts in which *kàban* is a clause-final particle too, albeit with the meaning ‘entirely, completely’. This concerns mainly the verbs *fɔ* ‘say’ and *dɔn* ‘know’ as predicates in negative sentences in the imperfective aspect.

- (30) *mánanin-ko dègun té fɔ kà bán án ká sigidaw lá.*
 plastic.bag-affair annoyance IPFV.AFF say completely 1PL GEN settlement PP
 ‘The annoyance through plastic bags in our neighborhoods is beyond words (lit. is not said completely).’ (kibaru548_04balo-mananinko_degun.dis.html)
- (31) *bì dènmisen té dɔn kà bán.*
 today child IPFV.AFF know completely
 ‘One does not completely know (understand) today’s youths.’ (tarawele-masalabolo_repl.html #13100929)

The reduced form when compared with the consecutive construction from which it originates, the restricted affirmative perfective context in which the PhP meaning is obtained and the negative isolation test are arguments in favor of considering *kàban* ‘already’ to be a particle.

5.2.2. *bán* ‘not yet’

The NOT YET-expression *bán* may be derived from the ALREADY-expression *kàban* ‘already’. It illustrates a more advanced position on the grammaticalization scale than the latter.

- (32) *ń m’ ò mén, ń m’ ò yé bán / *kàban*
 1SG PFV.NEG AN hear 1SG PFV.NEG AN see not.yet
 ‘I haven’t heard, I haven’t seen it yet.’ (#900664 mariko-masunkulu.dis.html)

The lack of the infinitive marker *kà* as consecutive connective provides evidence of this marker’s advanced degree of decategorization in a twofold way. First, if it is supposed to originate from *kà bán*, it can be considered to have undergone formal reduction. Another argument supporting the view that *bán* is not a verb anymore when it expresses phasal polarity, is the general absence of verbs juxtaposed with other verbs without a connective morpheme, with the exception of the verbs *tága* ‘go’ and *nà* ‘come’ in V1-position¹⁰.

It is impossible to isolate *bán*. This proves that the item is not autonomous.

¹⁰ These two verbs are primarily directionals that have itive and ventive function respectively. But they have further functions, among others as “markers of text connectivity” (cf. Bourdin 2008). An example may be the use of the verb *nà* which connects events and is translated by means of ‘and then’, as in the following sentence, extracted from a story in which a woman announces that the man who marries her will be made a sacrifice at her father’s grave when the father dies. While the first verb, *nà* bears the perfective affirmative suffix, the second, *sà* ‘die’ occurs as bare verb.

à yé mùso nìn fúru. mùso fà nàna sà.
 3SG PFV.TR.AFF woman DEM marry woman father come-PFV.ITR.AFF die
 ‘He married this woman. (And then) the woman’s father died’. (Dumestre 1989: 84)

- (33) *Q: í y' ò yé wà? A: *àyi, bán*
 2SG PFV.TR.AFF AN see Q no not.yet
 'Have you seen it?' 'No, not yet.'

Bán does not show semantic integrity. Thus, it has the meaning of 'not yet' exclusively in negative perfective contexts. When it occurs in affirmative (34) and in negative other than perfective contexts (35) in the clause final position, it has a different function, which Dumestre (2011a: 83) designates as 'indication of impatience or anger'. Only in this second function, *bán* shows the pronunciation variants *bánni*, *báni*, *bá* (35).

- (34) *á y' ò láje bán !*
 2PL IMPER.PL AN watch PTC
 'So, look at it!' (#1297635,sangare-suruku_ni_konyomusow.dis.html)

- (35) *Đolo, í té wúli bá !*
 Ngolo 2SG IPFV.NEG get.up PTC
 'Ngolo, so wouldn't you get up?' (#687696,kamara-weleli.dis.html)

The formal reduction from *kà bán* to *bán*, the specialization in negative perfective contexts as PhP-expression as evidence for the absence of semantic integrity and the unlikelihood to be isolated or moved to another than the final position in a clause qualify *bán* clearly as a particle.

5.2.3. *kàkɔɔ* 'already'

As one of the most polyvalent items in the Bambara lexicon, *kɔɔ* can function as the verb 'get or be old' (36), as the predicate adjective or state verb 'old' (37a), the modifying adjective 'old' (37b), or the the noun 'elder sibling' (38), among others.

- (36) *ù kɔɔ-la kójugu.*
 3PL get.old-PFV.ITR.AFF very
 'They got (and hence, are) very old' (Touré 1999:76)

- (37) a. *sègin-w ká kɔɔ,* b. *sègin kɔɔ-w dòn*
 basket QUAL.AFF old basket old-PL COPID
 'The baskets are old, they are old baskets'

- (38) *ń kɔɔ, à bé dī?*
 1SG elder.sibling 3SG COPLOC how
 My elder sister / brother, how is life (lit. how is it)?

It also occurs as the ALREADY item, resembling the polyvalent item *xóxó* in Ewe (van Baar 1997: 292). More concretely, the form *kàkɔɔ* 'already' unequivocally points to its verbal origin, for it bears the infinitive morpheme *kà*. In this respect and in respect to its clause-final position, it resembles *kàban*. Like the latter, *kàkɔɔ* developed from a consecutive construction introduced by the infinitive morpheme and is likely to follow clauses with

a verbal or a nonverbal predicate. At the end of clauses with nonverbal predicates, which have present time reference, it has the meaning of ‘already’. When placed in verbal clauses in the imperfective aspect (39) or in a construction including a participle and the identification morpheme *dòn*, which have a resultative-stative reading (40), the interpretation of *kàkəɔ* is ambiguous between the meaning ‘since long’, ‘for a long time’, ‘previously’ on the one hand and ‘already’ on the other hand. Frequently occurring verbs in such clauses are *dòn* ‘(get to) know’, *yé* ‘see’ or *kàlan* ‘learn’. As a consecutive construction *kà kəɔ* refers to the entire preceding event stating that ‘it got old’ or, in other words, that it happened long time ago. Evidence for a grammaticalization process is the use of *kàkəɔ* in a clause with a nonverbal predicate (41) and with the verb *kəɔ*, without yielding a durative event (42). The grammaticalization process goes hand in hand with the change of the writing convention from *kà kəɔ* to *kàkəɔ*.

(39) *án b' à dòn kàkəɔ*
 1PL IPFV.AFF 3SG know already
 ‘We know it already / since long’

(40) *à dòn-nen dòn kàkəɔ*
 3SG know-PTCP COPID already
 ‘It is already known / known since long’.

(41) *mùso mín-nu bé jí tà ù kùn ná kà à səɔ kán-dimi*
 woman REL-PL IPFV.AFF water take 3PL head PP INF 3SG find neck-ache
 ‘Women who carry water on their head although

bé ù lá kàkəɔ, kədimi bé sé kà òlu
 COPLOC 3SG PP ALREADY back-ache IPFV.AFF be.able INF 3SG.EMPH
 they have already neck pain, are likely to get back pain, as well.’

fána mìnɛ.

also catch

(dogotoro_13ka_caya.dis.html #4994697)

(42) *ù kəɔlən dòn kàkəɔ*
 3SG get.old-PART.PFV COPID since.long/already
 ‘They are already old / since long.’

Kàkəɔ cannot be removed from the final position, for instance, to the clause initial position. Keeping some of the properties of a consecutive construction, it is accepted exclusively in final position. As ‘already’, *kàkəɔ* is found mostly in affirmative clauses. However, negative polarity does not change its meaning, as illustrated in (43), a clause which bears the negative predicate marker *mán*. Because of its semantic integrity, *kàkəɔ* can be categorized as an adverb.

(43) *[məɔ kénɛman fána bé sé kà dòn jí súma ná wáati kùnkurunnin kà à kò]*
nkà ò fàrati ká bòn məɔkəɔɔba mà wàlimà məɔ mìn
 but 3SG danger QUAL.AFF big old.person PP or person REL

mán *kéne* *kàkɔrɔ*.
 QUAL.NEG healthy already

[Healthy people can go into cold water for a while / spend a while in cold water],
 ‘but this is very dangerous for old persons or for people who are already not
 healthy.’ (#4952594 dogotoro_10furakeli_kunfolo.dis.html)

5.2.4. *fɔlɔ* ‘not yet’

Fɔlɔ is another polyvalent item in Bambara (Dumestre 2011b: 5, Vydrin 2019: 371): it is a verb meaning ‘start with/by, do first’ (44), a noun meaning ‘former times, the past, the first one’ (45), a determiner (Dumestre 2011a: 339) or adjective (Vydrin 2019: 371; Bamadaba) meaning ‘the first X’ (46), an adverb meaning ‘in former times’ (Vydrin 2019:371; Bamadaba) (47), (48). As NOT YET PhP expression, illustrated in (17), it is subject to controversy, for it is classified as being “close to particles” by Dumestre (2011b: 5) and as adverb by Vydrin (2019: 371) and Bailleul (2007).

(44) *fɔlɔ* as a verb

ù *fɔlɔ-la* *kà* *zú* *siri* *ntúra-kɔrɔ-ba* *kán* *ná*
 they do.first-PFV.ITR.AFF INF joke attach ox-old-AUG neck PP
 ‘They attached the joke first to the neck of the big ox.’ (bamanankan_kalangafe1995-zup.repl.htm)

(45) *fɔlɔ* as a noun

nógɔndan *ó* *nógɔndan* *Ana* *bé* *ké* *fɔlɔ* *yé*.
 competition DIST competition PROP.N. IPFV.AFF become first PP
 ‘At every competition, Ana becomes first.’ (bamanankan_kalangafe1995-zup.repl.html)

(46) *fɔlɔ* as an adjective

túru-li *sàn* *fɔlɔ*, *mògɔ* *dó-w* *yé* *táari* *tán* *túru*.
 plant-NMLZ year first person certain.PL PFV.TR.AFF hectare ten plant
 [We started to plant in 1986, 1987.]‘The first year of planting, some people planted ten hectares’. (traore-hine_nanaIII.07.dis.html)

Applying van Baar’s criteria I will argue that *fɔlɔ* is an adverb in its meaning of ‘in former times’, but a particle when it has the meaning of ‘first’ and when it is a PhP expression.

Vydrin (1999: 89, 2019: 371) shows convincingly that some of the polyvalent items have a HL tune when they are used as adverbs, while they have a HH tune when they belong to the classes of nouns, verbs or numerals. Although the author does not list *fɔlɔ* ‘in the past’ among the lexemes that show this modified tune, my language assistant makes a clear distinction between the noun *fɔlɔ* ‘the first one’, realized HH on both syllables [fɔ́lɔ́], and *fɔlɔ* [fɔ́lò] ‘in the past’ realized HL. Thus, *fɔlɔ* [fɔ́lò] ‘in the past’ can be grouped together with the other adverbs, showing the characteristic tune.

ADV		other		word class
<i>kúnùn</i>	‘yesterday’	<i>kúnun</i> (‘) [kúnù]	‘yesterday’	N
<i>sísàn</i>	‘now’	<i>sísan</i> (‘) [sísá]	‘now’	N
<i>kélèn</i>	‘already’	<i>kélen</i> [kélĕ]	‘one’	NUM
<i>fǒlò</i>	‘in the past’	<i>fǒlò</i> (‘) [fǒlò]	‘(do) first, not yet’	ADJ, PTC, V

Table 5. Tonal distinctions of the polysemous items

The use of *fǒlò* as adverb ‘in the past’ is illustrated by the following examples:

- (47) *fǒlò, án tùn bé táa jí tà kòlòn ná.*
 in.the.past 1PL INAC IPFV.AFF go water take well PP
 ‘In the past, we went to fetch water at the well.’
- (48) *án tùn bé táa jí tà kòlòn ná fǒlò,*
 1PL INAC IPFV.AFF go water take well PP in.the.past
 ‘In the past, we went to fetch water at the well.’
sísan, án t’ ò ké bilen.
 now 1PL IPFV.NEG AN do anymore
 but now, we don’t do it anymore.’

Fǒlò ‘in the past’ occurs in different positions – clause initially and clause finally. Some speakers accept the clause-final position only when something follows (48). Further, as a response to a question, it can be isolated; however frequently in its reduplicated form.

- (49) *Q: ò tùn bé yèn wà? A: òwò. fǒlò ~ fǒlòfǒlò*
 AN INAC COPLOC there Q yes in.the.past in.the.past.REDUP
 ‘Did it exist?’ ‘Yes. In the past.’

The adverb *fǒlò* ‘in the past’ maintains its meaning under negation.

- (50) *fǒlò, án tùn té táa jí tà kòlòn ná.*
 in.the.past 1PL INAC IPFV.NEG go water take well PP
 ‘In the past, we didn’t go to fetch water at the well.’

The tune characteristic of adverbs as a formal criterion, the autonomy evident in a variable position at the front or at the end of a sentence and the identity of meaning in affirmative and negative context provide evidence that *fǒlò* ‘in the past’ is an adverb.

As PhP item, *fǒlò* ‘not yet’ is a particle. It is realized HH; it is not isolatable (51), and the only position possible is at the end of a clause (52).

- (51) *Q: à nà-na wà? A: *àyi, fǒlò*
 3SG come-PFV.ITR.AFF Q no not.yet
 ‘Has he come?’ ‘No, not yet.’

- (52) [Ntólofoori bé yàn dé ?]
 àyi. à má nà fɔ́lɔ . *fɔ́lɔ à má nà
 no 3SG PFV.NEG come yet
 ‘[Say, is N. here ?] No, he has not come yet.’ (dukure-ni_san_cyenna.dis.html)

The meaning of *fɔ́lɔ* changes under the influence of polarity: in an affirmative clause, for instance in the imperative (53) or in the future (54), it expresses ‘first’ and ‘from now on’; only in a negative clause in the perfective aspect, it has the function of the PhP expression ‘not yet’ (52).

- (53) nà yàn fɔ́lɔ!
 come here first
 ‘Come here first (for the moment)’. (Dumestre 2003:330 ; 2011a:339)

- (54) án ná à tò ké dògo lá fɔ́lɔ.
 IPL FUT.ASS.AFF 3SG rest do secret PP first
 ‘We will do it secretly from now on / at present.’ (Dumestre 2003:330)

The unlikelihood to be moved to another than the clause-final position and the restriction of the meaning as a PhP to sentences with negative polarity qualify *fɔ́lɔ* ‘not yet’ as particle, while the tune, the positional flexibility of *fɔ́lɔ* [fɔ́lɔ] ‘in the past’ and the maintenance of its meaning in both affirmative and negative sentences suit its categorization as adverb.

While in the cases of *kàban* and *bán* a certain degree of formal reduction pointed to grammaticalization, the absence of any bleaching of *fɔ́lɔ* justifies it to be a case of semantic specialization but not yet an advanced degree of grammaticalization. The French translation of ‘not yet’ as *d’abord* ‘first’ by some speakers suggests a close relation between those two meanings. Thus, *il n’est pas venu d’abord* ‘he has not come first’ can be heard for standard French *il n’est pas encore venu* ‘he has not come yet’. This context dependent semantic specialization can be interpreted as a preliminary step in the process of grammaticalization: under negation a non-negative non-PhP-item develops into a negative PhP expression (cf. van Baar 1997: 250, 253). In other contexts, *fɔ́lɔ* has other meanings.

The tonal similarity between the adjective ‘first’, the clause-final ‘first’ in affirmative clauses and ‘not yet’ in negative clauses suggests that the NOT YET-expression evolved from the adjective passing probably through the affirmative usage. All three, the ‘first’ as in *à mùso fɔ́lɔ* ‘his first wife’, ‘first, for the moment being’ and ‘not yet’ meanings are uttered from a perspective taking something else or a future phase into account. However, only in negative sentences where it translates by NOT YET *fɔ́lɔ* indicates that a situation does not hold at reference time relating it to a subsequent moment at which it holds.

5.2.5. *túguni* ‘still’ and ‘no longer’

The STILL and NO LONGER expression *túguni* (and its allomorphs *túgun*, *tún*) has its origin in the same word stating iterativity, which is represented by the following axis:

Van Baar (1997: 108) identifies two meaning components of iterativity. The first is the repetition (copy, represented by the second + in fig. 2) of an earlier situation (the first +) illustrated by the sentence *John came yesterday; today he came again*. The second

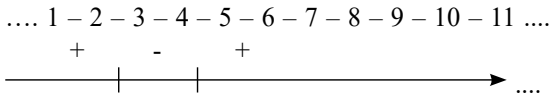


Fig. 2: Iterativity (according to van Baar 1997: 108)

one is the return of a situation to the original, as in *Shortly after Peter had pulled up the heavy weight, he had to release it again*. In Bambara, *túgun* covers both components of the semantics of iterativity. In (55), the fish that spoke before Fatumata cooked it, talks again after having been cooked. The fact of speaking can be understood as a copy of the previous uttering. In (56), the liberation of the chicken is the return of the original situation, before it was caught, i.e. when it was free.

- (55) à y' i kánto túguni
 3SG PFV.TR.AFF REFL.2SG speak again
 [Fatumata caught a fish, gutted it, cut it in pieces. Before she cooked it, the fish spoke to her. Then, the fish was cooked], 'it spoke again.' (179338 bamanankan_kal-angafe1995-zup.repl.html)

- (56) à yé syè mìnè k' à bila túgun
 3SG PFV.TR.AFF chicken catch INF 3SG let again
 'He caught the chicken and let it free again.'

The two iterative meanings are also attested in negative sentences, in (57) the return-to-the-original meaning, in (58) the copy-of-an-earlier-situation meaning.

- (57) i bólo fila kórɔ̀talen bé cogo mín,
 2SG arm two raise-PART COP.LOC.AFF way REL
 à ní bére té jìgin túgun dé!
 3SG and stick IPFV.NEG lower again PTC

[Da was about to smash a cane on Bakari Dian's head, Simbalan appeared and prevented him from doing this:] "The way you hold your two arms, you will not lower the one holding the cane again" (Dumestre 1979: 100)

- (58) jínè-ke túnun-na. kàbini ò ké-ra,
 jinn-man get.lost-PFV.ITR.AFF since an happen.PFV.ITR.AFF

m̀ɔ̀gɔ sí má jínè yé túgun.
 person no PFV.NEG jinn see again

[the jinn came each rainy season to the village.... then:] 'it disappeared. Since this happened, nobody saw the jinn again.' (#104254,bird_hutchison_kante-an_ka_bamanankan.dis.html)

In sentences that are construed as imperfectives, progressives or states, *túguni* has the meaning of 'still' in affirmative clauses, and 'no longer' in negative clauses. Compared

with the iterative which makes reference to both phases at which the event holds – the original one and its copy or return-to-its-origin state, the NO LONGER meaning of *tún* focuses only on the negative phase and states the non-occurrence of an event with reference to a preceding phase at which it holds. *túgun* in the negative sentence (59) states that the situation does not hold making reference to a previous situation at which the addressee was working. The cattle’s activity of chewing grass in (60) does of course include pauses in which it does not chew, but the imperfective suggests the continuous activity to have stopped after a period where it took place.

Túguni in the sense of NO LONGER occurs in locative clauses, as in (59).

- (59) state reading
dùguba sigi ká gèlen, é dùn té báara lá túgun.
 city life QUAL.AFF difficult 2SG.EMPH PTC COP.NEG work PP no.longer
- án té táa fà-so lá ?*
 1PL IPFV.NEG go father-house PP
 ‘Life in a city is expensive, you do not work (lit.: are not at work) anymore, shouldn’t we go back home?’ (269530 bamanankan_kalangafe5-zup.repl.html)

The schematic representation of NO LONGER is as follows:

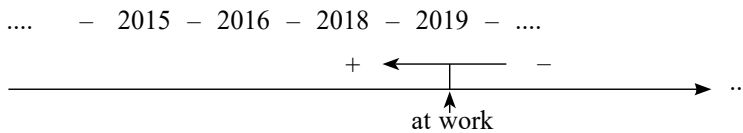


Fig. 3: The neutral scenario of NO LONGER

Túguni in the sense of ‘no longer’ is used in imperfective construction.

- (60) *à té ná-nimi-ni ké túguni.*
 3SG IPFV.NEG grass-chew-NMLZ do no.longer
 ‘[The cattle has fever, it trembles ...] It does no longer chew/ruminate grass’ (#150972 baganw_ka_banaw-zup.old.repl.html)

However, affirmative examples with a STILL meaning are very rare (61), and my consultants do not find them acceptable. Van Baar (1997: 191) observed a similar distribution of a NO LONGER / STILL item originating from an iterative in Tongan and Ewe. In Bambara both *túguni* and *bilen* show a comparable distribution.

Túguni in the sense of STILL:

- (61) *nkà ñ t’ à dón k’ à fɔ à bé bálo lá tún*
 but 1SG IPFV.NEG 3SG know INF 3SG say she COP.LOC.AFF life PP still
 [...] but I don’t know if she is still alive (banbera-faamanje_ni_faantanje.dis.html; 13145602)

As marker of iterativity, *túguni* can be isolated with some reservations. For instance, it can be repeated after the interjection *ɔwɔ* ‘yes’. However, it is not likely to be moved to clause-initial position.

- (62) *à nà-na túguni, ɔwɔ, túguni.*
 3SG come-PFV.ITR.AFF again yes again
 ‘He came (back) again, yes, again.’

**túguni, à nà-na.*

In its function as marker of one of the two senses of iterativity, *túguni* is stable under negation. The comparison of (55) and (58) provides evidence for the copy-sense of the iterative, the comparison of (56) and (57) evidence for the return-to-the-original-sense of the iterative; both meanings do not change under the influence of polarity. However, as PhP expressions the meaning of *túguni* is highly dependent on the context: as STILL expression it shows that a state holds at reference time relating it to a subsequent period where it doesn’t hold, as NO LONGER it expresses that the state does not hold at reference time relating it to a previous period where it did.

Therefore, while iterative *túguni* appears to be closer to adverbs, the lack of autonomy and of semantic integrity are arguments to categorize the PhP item *túguni* rather as a particle. In contexts expressing states and continuous events, *túguni* has been specialized as PhP expressions. The formal reduction of *túguni* to *túgun* and *tún*, however, cannot be attributed to the grammaticalization process from the adverb with iterative meaning to PhP items, for the shortest allomorph *tún* occurs in both senses, the iterative and the PhP.

5.2.6. *bilen* ‘still’ and ‘no longer’

bilen seems to be a dialectal variant of *túguni* in both its functions as iterative ‘again’ and the PhP expressions STILL and NO LONGER. *bilen* belongs to the Bambara dialect of Segú; it occurs marginally in affirmative sentences with the meaning of iterative ‘again’ (63) and of the PhP expression STILL (64), uses that are nevertheless rejected by speakers from Bamako.

- (63) *ń í nà-na bilen, í ń’ à dón*
 if 2SG come-PFV.ITR.AFF again 2SG FUT 3SG know
 ‘If you come again, you will know it.’ (French : ‘Si tu reviens, tu va voir !’ ; Bama-daba, sentence illustrating *bilen*)

- (64) *ò fàna ní sísan cé, né má sé k’ à báara dábila.*
 an PTC and now between 1SG PFV.NEG be.able INF 3SG work abandon
 ‘Since that time, I was not able to abandon this work.’

ń bé ò báara lá bilen
 1SG COP.LOC.AFF AN work PP still

‘I am still doing that work (lit. I am still at that work).’ (basiya.dis.html #76987)

More characteristic of *bilen* is its negative value (Dumestre 2003: 31), and in negative sentences it is accepted even by speakers from Bamako who do not use it actively themselves. (65) illustrates its iterative function in a negative clause, (66) its function as the PhP expression NO LONGER.

(65) *sàma má wúli bilen.*
 Elephant PFV.NEG get.up again
 [When Elephant said the forbidden word, he immediately fell to the ground ...]
 ‘Elephant did not get up again.’ (sonsannin_surukuba.dis.html #7320202)

(66) *í mán júgu bilen.*
 2SG QUAL.NEG evil no.longer
 ‘You are not evil anymore’ (Dumestre 2011a: 113)

The distribution of *bilen* resembles that of *túguni*: with dynamic verbs used in the perfective aspect, it has iterative meaning regardless of the polarity the clause bears. In affirmative clauses construed as states or continuous events, *bilen* can be considered as a PhP expression – namely as a STILL-expression in affirmative clauses and as a NO LONGER-expression in negative clauses. In none of its uses is *bilen* likely to be isolated¹¹. Therefore, the criterion of semantic integrity of *bilen* as iterative allows to categorize it as adverb; on the contrary, due to the changing meaning of *bilen* as PhP item under negation in states and continuous events¹², the PhP item is categorized as particle.

5.2.7. *háli bì, háliisa*

Háli bì or *háli bì*, realized [hálibì] or [hálibì], and *háliisa* with its tonal realization [háliisa] or [háliisa] are compound forms containing the element *háli*, expressing continuation and inclusion and translated by ‘even’, ‘until’ and an element signaling inclusion, either *bì* ‘today’ or *sá*. Dumestre considers *sá* to be originally a short form of *sísan* ‘now’, which is however hardly used as an adverb anymore¹³. As an independent morpheme, *sá* is nowadays

¹¹ Although *bilen* occurs in the position immediately following the subject, its function differs considerably from the same element used in clause-final position. In post-subject position it has the function of ‘however’ and of a negative hypothetical predicate marker, nowadays fallen into disuse.

¹² In the following sentence in the future in which *bilen* has the sense of NO LONGER there is the idea of a state in which the person is in a risk of death. ‘He will no longer die’ is used in the sense of ‘the injured person is not in a risk of death anymore’.

ní sà yé màa mîn cín, ní ò nà-na à káanamugu
 if snake PFV.TR.AFF person REL bite if this come-PFV.ITR.AFF 3SG magic.powder
mîne k’ à dún, ò tîgi be kénéya, ò te sà bilen.
 take INF 3SG eat this person IPFV.AFF recover this IPFV.NEG die no.longer
 ‘If someone bit by a snake takes the magic powder, s/he recovers, s/he is no longer in the risk of death’. (dumestre-manigances_2002_12.dis.html #197299)

¹³ For Dumestre, *sísan* ‘now’ is a noun, for it is able to fill the syntactic role usually filled by nouns. I concur with Vydrin (1999: 89) who considers the nominal use of *sísan* as a case of conversion. An argument corroborating this view is a different tune that distinguishes its adverbial use from its nominal use. Thus, *sísan* is realized [sísàn] as an adverb and [sísán] as a noun.

a discourse particle that conveys less a temporal meaning than a sense of insistence or irritation. This may be the reason why the convention distinguishes to write *háli sa* as one word and *háli bi* as two words, signaling the original and still existent independence of each item of the latter expression. The maintenance of the separate writing for *háli bi* points to speakers' awareness of the composite meaning of *háli bi*, in which the sense of *bi* 'today' is sometimes extended to 'the present'. While *bi* and *sa* are items of Mande origin, *háli*, occurring in both STILL expressions, comes from Arabic. I have noted elsewhere (Dombrowsky-Hahn to appear) hypotheses on the Arabic item of origin.

- (67) *dén-w bé súnɔgɔ lá háli bi / háli sa.*
 child-PL COP.LOC sleep PP still
 [Talking in the afternoon about the children who took a nap after lunch.] 'The children are still sleeping.'

The inspection of the criteria of autonomy and semantic integrity, decisive of the items' status as adverbs or particles, shows that *háli bi* and *háli sa* are adverbs: they can alternatively be moved to sentence-initial position (68), and it is possible to isolate them (69).

- (68) *háli bi / háli sa déw bé súnɔgɔ lá*
 still child-PL COP.LOC sleep PP

- (69) A: *í bé mún ná?*
 2SG COP.LOC what PP
 'What are you doing?'

B: *ń bé tóbi-li lá.*
 1SG COP.LOC cook-NMZR PP
 'I am cooking'

A: *háli bi?*
 still
 'Still?'

B: *ɔwɔ, háli bi.*
 yes, still
 'Yes, (I am) still (cooking).'

5.2.8. *kélèn*

Kélèn is used in counterfactual contexts as a marked version of ALREADY. Many attested examples are uttered in contexts where a guest is about to leave at a moment considered premature by the host.

- (70) A: *à y' í sára k' à béna táa dé.*
 3SG PFV.TR.AFF REFL announce that 3SG FUT go DP
 'He says he is going to leave'

B: à b́éna táa kélèn?
 3SG FUT go already
 ‘He is going to leave already?’

There is no unanimity among the speakers whether it is possible to replace *kélèn* by the unmarked ALREADY expression *kàban*. For instance, not all of them accept *kàban* in a clause with future time reference.

(71) (*) à b́éna táa kàban?

Kélèn seems to be related to the numeral *kélen* [kélé] ‘one’, and the quantitative ‘alone’, ‘only’; however it is not clear, how this is related to ‘already’. It is only the HL tune that distinguishes it from the numeral, which has HH tune.

The HL tonality which *kélèn* shares with a number of other adverbs and the autonomy criterion illustrated in the dialogue extract in (72) allow the categorization of *kélèn* as an adverb.

(72) Paul: wáati sé-ra, fòli dòn.
 time arrive-PFV.ITR.AFF greeting copid
 ‘The time (for me to leave) has come; it’s just a hello’

Jigi: kélèn?
 already
 ‘Already?’ (Morales et al. 1991: I,30)

6. Summary: the assignment of PhP-items to word classes revisited

The linguist glossing sentences takes at the same time a decision about the assignment of individual items to a particular word class. Whether the Bambara PhP term *túgumi*, for instance, is glossed as “again” or “not yet” or “still” is also a decision about its categorization. In the previous sections I studied the categorization of the Bambara PhP expressions ALREADY, NOT YET, NO LONGER AND STILL using van Baar’s (1997) insights from a crosslinguistic study and his theoretical approach to the topic. The results can be resumed as follows. First, PhP expressions occur (mainly) utterance-finally, which is the position shared by adverbs and particles in Bambara. Their position thus suggests classifying them as one or the other. Although adverbs and particles resemble each other also in other respects, there are important differences between them, too. Adverbs are considered to be modifiers of non-nominal heads. They are isolatable from the rest of the sentence and are autonomous with regard to their meaning, which for instance does not change under the influence of negation. By contrast, particles are linkers that connect a part of a sentence to a wider textual or extra-textual context, from which they cannot be separated. They do not have autonomous meaning, which is sensitive to negation. Particles are highly grammaticalized elements which have undergone formal reduction with respect to their sources, that have

lost other properties of their sources and whose function depends on the context of their occurrence. There is a variety of lexical items (nouns, verbs, adjectives) which can be the source of the grammaticalization process resulting in a particle, among others adverbs.

Van Baar (1997) suggests three principles to distinguish adverbs and particles: the likeliness of particles to be collocated with nominal heads without showing agreement features, the autonomy principle and the criterion of semantic integrity. Their application to the Bambara items used as PhP expressions has led to their recategorization, resumed in table 7. Most of them are categorized as adverbs, an adverbial phrase or particles. Some PhP items (*kàban*, *túguni*, *bilen*, *fóló*) coexist with their source elements without having undergone formal reduction. Where there is no formal distinction say between an adverb and a particle that developed from it, it is especially their autonomy (likeliness to be moved to sentence-initial position) and their semantic integrity, i.e. the sensitivity to negation, that distinguishes the PhP expression from the respective source element.

Thus, while many instances of multiple categorization of elements in the Bambara lexicon is due to the process of conversion of one major word class (noun, verb or adjective) to another, others should be recognized as the result of (an ongoing process of) grammaticalization, including the change from an adverb being decategorized to a particle.

PhP item	expression	source	reconsidered	other
ALREADY	kàban	INF+V	PTC	kà (à) bân: consec. constr. 'finish'v
NOT YET	bán	PTC	PTC	
ALREADY	kélèn		ADV	kélèn 'one' NUM
NO LONGER	túguni	ADV	PTC	iterative: ADV
STILL	túguni	NEG.PTC	PTC	
STILL	háli bì	PTC+N	ADV phrase	
STILL	hálisa	PTC+N	ADV	
NO LONGER	bilen	ADV	PTC	iterative: ADV
STILL	bilen	NEG.PTC	PTC	
NOT YET	fóló [fólól]	ADJ	PTC	[fólól] 'first; [fólò] 'in the past' ADV
ALREADY	kàkɔɔ	INF+V	ADV	kà kɔɔ consec. constr.

Table 6: Categorization of PhP expressions reconsidered

Abbreviations

1,2,3	1 st , 2 nd , 3 rd person	COPEQU	copula in equational clauses
ADJ	adjective	COPID	copula in affirmative presentative clause
ADV	adverb, adverbial	COPLC	copula in affirmative locative clause
AFF	affirmative	D	demonstrative
AN	anaphoric pronoun	FUT	future affirmative
ART	article	FUT.ASS.AFF	future assertive affirmative
AUX	auxiliary	FUT.NEG	future negative
COMP.N.	compound noun		
COP.NEG	copula negative		

GEN	connective between possessee and possessed in genitive construction	PP	postposition
IMPER	imperative	PTC	particle
INAC	inactuality marker	PTCP	participle
INF	infinitive	Q	yes/no question particle
INTERJ	interjection	QUAL.AFF	qualitative affirmative
IPFV.AFF	imperfective affirmative	QUAL.NEG	qualitative negative
IPFV.NEG	imperfective negative	REDUP	reduplication
N	noun	REFL	reflexive
NMLZ	nominalization	REL	relativizer
PFV.I.TR.AFF	perfective intransitive affirmative	SBJV.AFF	subjunctive affirmative
PFV.NEG	perfective negative	SBJV.NEG	subjunctive negative
PFV.TR.AFF	perfective transitive affirmative	SG	singular
PL	plural	v	verb

References

- Baar, Tim van. 1996. "Particles". In: Betty Devriendt, Louis Goossens & Johan van der Auwera (eds.). *Complex structures. A functionalist perspective*. Berlin: De Gruyter Mouton, 259-301.
- Baar, Tim van. 1997. *Phasal polarity*. Dordrecht: Foris Publications.
- Bailleul, Charles. 2007. *Dictionnaire bambara-français. Troisième édition corrigée*. Bamako: Donniya.
- Bailleul, Charles, Artem Davydov, Anna Erman, Kirill Maslinsky, Jean J. Méric & Valentin Vydrin. 2011-2018. *Bamadaba: Dictionnaire électronique bambara-français, avec un index français-bambara*. <http://cormand.huma-num.fr/bamadaba.html>.
- Bourdin, Philippe. 2008. "On the grammaticalization of 'come' and 'go' into markers of textual connectivity". In: Maria López-Couso & Elena Seoane (eds.). *Rethinking grammaticalization. New perspectives*. Amsterdam: Benjamins, 37-59.
- Creissels, Denis. 1985. "Les verbes statifs dans les parlers manding". *Mandenkan* 10,1-32.
- Dombrowsky-Hahn, Klaudia. 2012. "Motion events in Bambara (Mande)". *Journal of African Languages and Linguistics (JALL)* 33(1), 37-66.
- Dombrowsky-Hahn, Klaudia (to appear). "Phasal polarity expressions in Bambara (Mande): pragmatic distinctions and semantics". In: Raija Kramer (ed.). *The expression of phasal polarity in Sub-Saharan African languages*. Berlin: De Gruyter Mouton.
- Dumestre, Gérard. 1979. *La geste de Ségou racontée par des griots bambara*. Paris: Armand Colin.
- Dumestre, Gérard. 1987. *Le bambara du Mali: essais de description linguistique. Thèse de Doctorat d'Etat*. Directeur de thèse: Maurice Houis. Paris: Université Paris III et INALCO.
- Dumestre, Gérard. 1989. *La pierre barbue et autres contes du Mali. Edition bilingue bambara – français*. Angers: Ville d'Angers, Bibliothèque Municipale.
- Dumestre, Gérard. 1999. "Des auxiliaires en bambara". *Mandenkan* 35,1-16.
- Dumestre, Gérard. 2003. *Grammaire fondamentale du bambara*. Paris: Karthala.
- Dumestre, Gérard. 2011a. *Dictionnaire bambara – français suivi d'un index abrégé français – bambara*. Paris: Karthala.
- Dumestre, Gérard. 2011b. "À propos des adverbes du bambara, ou de l'art d'accommoder les restes". *Mandenkan* 47,3-11.
- Heine, Bernd. 1993. *Auxiliaries: Cognitive forces and grammaticalization*. New York: Oxford University Press.
- Idiatov, Dmitry. 2000. "Le sémantisme des marqueurs aspecto-temporels du bambara: une tentative d'analyse". *Mandenkan* 36,1-59.
- Kone, Kassim G. 2010. *bamanankan danyegafe*. 2nd edition (first edition 1995). West Newbury, Massachusetts: Mother Tongue Editions.
- Lehmann, Christian. 1982. *Thoughts on grammaticalization. (Arbeiten des Kölner Universalien-Projekts 48)*. Cologne: Universalien-Projekt.

- Morales, José, Casimir Sinayogo & Charles Teisserenc. 1991. *Baro ka di bamanankan na. Parlons bambara avec plaisir. Fascicule I*. Falaje: CEL.
- Touré, Mohamed. 1999. *Bambara Lesebuch. Originaltexte mit deutscher und französischer Übersetzung. Livre de lecture Bambara. Textes originaux Bambara avec traductions allemandes et françaises. Mit Zeichnungen von Melanie Leucht*. Cologne: Köppe.
- Vydrin, Valentin. 2019. *Cours de grammaire bambara. Ouvrage en réalité augmentée*. Paris: Presses de l'INALCO.
- Vydrin, Valentin. 1999. "Les parties du discours en bambara: Un essai de bilan". *Mandenkan* 35,73-93.
- Vydrin, Valentin, Kirill Maslinsky, Jean Jacques Méric & Andrij Rovenchak. (2011–2018). *Corpus Bambara de Référence*. <http://cormand.huma-num.fr/>

Ma-causatives in Tomo Kan Dogon: between causatives and passives

Abstract

This article deals with the causative/passive polysemy in the Tomo Kan language (< Dogon < Niger-Congo). The aim of the article is to describe semantic and syntactic properties of the distant causative *-má* and the homonymous modal passive which are instantiations of the same morpheme. First, I show that the causative can be analyzed as a marker of distant causal relation, whereas the modal passive is a marker with the meaning of participant-external possibility. Second, I discuss a possible polysemy scenario based on the structural properties of the causative. Namely, I will argue that the passive reading could not have arisen from permissive readings of the causative marker and it is better to describe the shift in terms of subject deletion, which led to the reanalysis of the syntactic structure.

Keywords: verb morphology, valency changing alternation, causative, passive, Niger-Congo languages, Dogon languages, Tomo Kan language

1. Introduction

The causative/passive polysemy is a well-known property of several language families and it has been the topic of a large discussion up to this day¹. However, the research has been primarily limited to descriptive work and no formal explanation was proposed for the phenomenon. One of the most persuasive attempts to account for this polysemy were some observations made by Haspelmath (1990). He argues that the causative-to-passive shift goes through the intermediate stage of reflexive causatives. The following examples illustrate the shift. In (1), a causative construction comprising the verb ‘have’ is given. If the causer at some point becomes co-indexed with the causee like in (2), it facilitates the identification of the two participants leading to a passive construction in (3).

- (1) *I have the barber shave me.* = ‘I CAUSED the barber to shave me’
- (2) *I have myself shaved by the barber.* = ‘I CAUSED MYSELF to be shaved by the barber’.
- (3) *I am shaved by the barber.*

¹ see, for example Haspelmath (1990) and Kulikov (2001) on the typology of the phenomenon, Kor-mushin (1976) on Turkic languages, Say (2013) on Kalmyk, Yap & Shoichi (2003), Rhee & Koo (2014) on Korean, among many others.

According to this scenario, a causative marker does not specify the causal relation, and intermediate events which led to the caused sub-event are not described by the verb. Under this reading, the relation between the agent's intention and the ongoing event is quite abstract and general. The next step to the passive meaning is the reanalysis of the agent as the cause of the event ('X made so that Y happened') and loss of agentivity. In (2), the sentential subject ('I') is not an individual controlling the action and, thus, it becomes a passivized agent, as in (3).

In this paper, I will try to show that this scenario does not seem plausible for the Tomo Kan language data. In section 2, I give a brief overview of Tomo Kan verb morphology. In section 3 I discuss the properties of *má*-causatives which gave rise to the passive marker. I describe semantics of modal passives in section 4, and show that the source of the modal meaning is the Futurum/Habitual suffix attached to these forms. The passive component of these constructions is contributed by the suffix *-má* and the verbs exhibiting this suffix lack a subject. In section 5 I show that considering the structural properties of causatives mentioned above, it is much more likely that the emergence of passive meanings is associated with the loss of the clausal subject rather than with the reinterpretation of transitive clauses.

2. Some features of the Tomo Kan morphology

Tomo Kan is a Dogon language spoken by approximately 180 000 people in Mali and Burkina-Faso. The Tomo Kan described here is spoken in the village of Ségué. This dialect was described by Léger (1971) and Dyachkov (ms.).

Tomo Kan is a verb-final SOV language, but the order of constituents may vary depending on the information structure. Like most African languages, Tomo Kan is a tonal language and many syntactic processes can be accompanied by tonal changes. The verbal morphology is quite rich and comprises several derivational affixes (causatives and reversives), aspect markers and negation markers. Person agreement is almost absent, with the exception of 3PL forms which agree with the subject. A typical verbal form consists of the following slots:

- (4) stem – causative/reversive – aspect – negation – past time

3. *má*-causatives

má-causatives are productive causatives which can be derived from almost any stem including stems with other unproductive causative suffixes (which are *-rV-* and *-gV-* with underspecified vowels). The latter can be described as contact causatives while the *má*-causatives are distant causatives. These two types are exemplified in (5) and (6), respectively.

- (5) ηkò ^{LH}sùbá mí: àbí-rí
 DEM mat 1SG.EMPH be.spread-CAUS.PFV
 'It was me who spread the mat'.

- (6) ɲkò ^{LH}sùbá mí: sè:dú
 DEM mat 1SG.EMPH S.
 àbà-rà-mí
 be.spread-CAUS-CAUS.D.PFV
 ‘It was me who made Seydou spread the mat’.

In spite of the fact that distant semantics is not always implied in clauses with *má*-causatives, I call them distant/indirect causatives, since the meaning of distant causation is obligatory when the suffix is attached to minor causatives. In (5), for example, the agent is the only entity which causes the mat to be spread while in (6) there is an intermediate agent which is forced (or permitted) by another agent to spread the mat. Thus, the indirect impact on the mat is conditioned by the presence of the intermediate agent.

I assume that additional factors (for example, intermediate actors in examples given above) trigger the semantics of distant causation. However, descriptive properties of this causation are underspecified and the causative can express various kinds of causal relations. For instance, *má*-causatives may denote events which are brought to life without any specific intention of the external causer (=the agent). As a result, *má*-causatives are felicitous with adverbials meaning involitivity (8).

- (7) ɲ yìʔijé: kùgò ɲàʔà^L tà^L kúⁿ òbòrò-mí
 1SG unintentionally hen Guinea.fowl egg on sit-CAUS-CAUS.D.PFV
 ‘I accidentally put the chicken eggs under the Guinea fowl’.
- (8) ɲ jí: há jingà^L ló bònù-gò-mí
 1SG water REFL wish DAT be.troubled-CAUS-CAUS.D.PFV
 ‘I troubled the water intentionally’.

má-causatives must include a causee participant in their structure, even if the latter is not expressed overtly. The causee participant (i. e., the internal or intermediate agent) is marked with accusative case, if the causer is volitive:

- (9) ɲ sè:dù-lí íwⁿ dàngà-rà-mí
 1SG S.-ACC child sit-CAUS-CAUS.D.PFV
 ‘I told Seydou to seat the child’.

In Tomo Kan, the accusative case can mark only themes and is optional. Its uses are regulated by the animacy hierarchy (Comrie 1981; Aissen 2003) and the accusative is compatible mostly with NPs denoting animate or prominent referents (see Dyachkov ms. for details). However, the marking is optional in causative clauses as well, as is shown in (10). The position of the causee is fixed – it occupies the position before the internal VP so that the internal direct object is obligatorily preceded by the causee, see (11).

- (10) a. àmàdú sè:dù-lí kò ^Hnáʔná ɲú:-mí
 A. S.-ACC DEM cow eat-CAUS.D.PFV

b. àmàdú sè:dú kò ^Hnáʔná ɲú:-mí
 A. S. DEM cow eat-CAUS.D.PFV
 ‘Amadou made Seydou feed the cow’.

- (11) *à:mádú kò ^Hnáʔná sè:dù-lí ɲú:-mí
 A. DEM cow S.-ACC eat-CAUS.D.PFV
 Int.: ‘Amadou made Seydou feed the cow’.

Accusative marking is always a default option and oblique case markers are rejected (12). However, in other cases the Dative marker is possible. In contexts such as (13), the cow can be perceived as an auxiliary object by means of which the action is performed. All Dative-marked NPs in Tomo Kan are clause-level adjuncts denoting instruments, benefactives, materials, etc. Accusative marking is impossible in this case and this is the reason why the NP denoting the causee cannot be analyzed as a core argument.

- (12) *sàtá sá íw ^Lò ɲá: jùmbò-mà=:wó
 S. REFL child DAT fire fan-CAUS.D-IPFV
 ‘Sata made his child fan the fire’.

- (13) mí: náʔná lò/*-nì jí: bùnú-gò-mí
 1SG.EMPH cow DAT/ACC water be.troubled-CAUS-CAUS.D.PFV
 ‘I troubled the water by a/the cow (e.g., by having driven it into the water)’.

The causee can be easily omitted from the sentence. Since the causative encodes the semantics of indirect causation, the internal agent (=causee) is always preserved in the semantic structure. Therefore the following sentences lacking overt causees are possible:

- (14) ɲkò ^{LH}sùbá mí: àbà-rà-mí
 DEM mat 1SG.EMPH be.spread-CAUS-CAUS.D.PFV
 ‘It was me who spread the mat [by making somebody do it]’.

- (15) sè:dú ɲkò ^{LH}sùbá àbà-rà-mí
 S. DEM mat be.spread-CAUS-CAUS.D.PFV
 ‘It was Seydou who spread the mat [by making somebody do it]’.

Further evidence for the presence of the internal agent in the structure of *má*-causatives can be obtained if one compares them to contact causatives. Semantically, *má*-causatives can be easily distinguished from this set of causatives since they are incompatible with adverbials like *há hinnà* ‘by oneself’. However, this adverbial is felicitous with contact causatives signaling that only *má*-causatives require an obligatory intermediate agent:

- (16) ɲ [↑]ɲkó ^{LH}sùbá há hinnà àbí-rí
 1SG DEM mat REFL one be.spread-CAUS.PFV
 ‘I spread the mat myself’.

- (17) *ɲkò^{LH} sùbá mí=: há hinnà àbà-rà-mí
 DEM mat 1SG.EMPH REFL one be.spread-CAUS-CAUS.D.PFV
 ‘I spread the mat myself’.

Syntactically, *má*-causatives are monoclausal. The causative verb denotes a single event and it cannot be decomposed, either semantically and syntactically, into the causing sub-event (which is expressed by the suffix) and the sub-event brought to life by the causing event (=the caused sub-event). As a consequence, the ellipsis of only the causing subevent (expressed by the *-má* suffix) in (18) results in a semantically bizarre interpretation rejected by my consultant. Thus, in Tomo Kan caused sub-events alone cannot be taken into the scope of ellipsis and the only option is when causative verbs are present in both clauses (19).

- (18) *ɲ há sùng-ìⁿ bwèlè-mà=:wó
 1SG REFL rope-DIM pass-CAUS.D-IPFV
 sè:dú húnú yé bwèlè=:wó
 S. also there pass-CAUS.D-IPFV
 Int.: ‘I made my thread pass, and Seydou’s thread also passed’.

- (19) ɲ há sùng-ìⁿ bwèlè-mà=:wó
 1SG REFL rope.DIM pass-CAUS.D-IPFV
 sè:dú húnú yé bwèlè-mà=:wó
 S. also there pass-CAUS.D-IPFV
 ‘I passed my thread, and Seydou did so with his thread’.

There is also evidence that NPs denoting the causee (=internal agent) are not reconstructed when ellipsis takes the causative verb in its scope. In the case of (20), NPs denoting the theme and the causee are omitted in the second clause. However, the causee is not reconstructed in this case, suggesting that it is only the feeding event (‘X fed the cow’) that is performed by Sata but it is not Seydou who was the intermediate performer of the event. Semantically, it means that only the feeding event is affected by the ellipsis but not the causee. Once again, the only option to include the causee in the interpretation of a sentence is to repeat his/her name (21).

- (20) àmàdú sè:dù-lí kò ^Hnáʔná jú:-mí
 A. S.-ACC DEM cow eat-CAUS.D.PFV
 sátá húnú jú:-mí
 S. also eat-CAUS.D.PFV
 a. ‘Amadou fed the cow (by making Seydou do it), and Sata also fed the cow’.
 b. ‘*Amadou made Seydou feed the cow, and Sata made Seydou feed the cow’.

- (21) àmàdú sè:dù-lí kò ^Hnáʔná jú:-mí
 A. S.-ACC DEM cow eat-CAUS.D.PFV
 sátá húnú sè:dù-lí kò ^Hnáʔná jú:-mí
 S. also eat-CAUS.D.PFV DEM cow eat-CAUS.D.PFV
 ‘Amadou made Seydou feed the cow, and Sata made Seydou feed the cow’.

The two tests show that the caused sub-event cannot be separated from the causing event, that is, it does not constitute a separate clause.

Another argument for the monoclausal structure is the fact that temporal adverbials cannot modify two clauses separately. Crosslinguistically, biclausal indirect causatives can combine with two temporal adverbials having contradictory meanings, one of which modifies the caused sub-event and the other the causing sub-event, so that the two situations described are temporally non-adjacent. In Tomo Kan it is not possible, therefore the structure is monoclausal and the causative denotes a single event.

- (22) *ʔyá: minátá mí-nì ɲà: ʔyé: jòngò-mí
 yesterday M. 1SG-ACC food today mix-CAUS.D.PFV
 Int.: ‘Yesterday Minata made me mix the food today.’

Finally, evidence for monoclausality can be provided by reflexive binding. In Tomo Kan, reflexives can be bound by clausal subjects. In (23), the theme and the subject are co-indexed. In (24), the subject binds the cause and the internal theme at the same time.

- (23) sàtá há nàʔⁿà ɲú:-mí
 S. REFL cow eat-CAUS.PFV
 ‘Sata_i fed her_i cow’.
- (24) àmàdú há ɲǰáⁿ-nì há nàʔⁿà ɲú:-mí
 Amadou REFL younger.brother-ACC REFL cow eat-CAUS.PFV
 ‘Amadou_i made his_i brother feed his_i cow’.

However, the causee cannot bind reflexives. In (25), the reflexive can be only interpreted as bound by the subject (‘Amadou’s cow’) but not by the causee (‘elder brother’s cow’). If the internal direct object needs to be bound by the causee, then the simple (personal) pronoun but not reflexive is used. At the same time, no coindexation of the subject and the causee is possible, as shown in (26).

- (25) àmàdú há ɲǰáⁿ-nì_i wò_i ^{LH}nàʔⁿà ɲú:-mí
 Am. REFL elder.brother-ACC 3SG.POSS cow eat-CAUS.PFV
 ‘Amadou_i made his brother_i feed his_i cow’.
- (26) *sátá há-li kò ^Hnàʔⁿà ɲú:-mí
 S. REFL-ACC DEM cow eat-CAUS.PFV
 Int.: ‘Sata made herself feed the cow.’

The tests listed above lead us to several conclusions. First of all, “distant” *má*-causatives are monoclausal structures, where the causing and the caused sub-event cannot be modified independently. Second, the NP denoting the causee is detached from the verb, both semantically and syntactically. Third, the clausal subject (=causer) can bind each of the verbal arguments (including the causee), but the causee is structurally lower and cannot bind NPs itself. If the causee is a true accusative-marked direct object, then the explanation

is straightforward – no argument other than the one occupying the subject position is able to bind reflexives. In section 4, I will present a correct syntactic structure which would capture all of the facts listed above.

4. Passives

It is likely that *má*-causative are the diachronic source of two other Tomo Kan markers, modal/passive *-ma* and (probably) the Hortative suffix, which will not be discussed here in detail. The causative/passive polysemy is well-known in other languages including some Dogon languages, though the phenomenon in question has not received detailed explanation yet. It is probable that pro-drop favors the ellipsis of the subject, which leads to the reanalysis of causative clauses as passive ones. In Tomo Kan, pro-drop is indeed possible and can be found in causative clauses as well:

- (27) sè:dù-lí kò ^Hnáʔná nù:-m-è:
 S.-ACC NONH cow eat-CAUS-PFV.3PL
 ‘Seydou was authorized to feed the cow’.

Modal passive forms are attested only with the Futurum/Habitual suffix. Other aspectual affixes, which may follow the *má* suffix were not attested, and forms like *ʔwà:-mà-wó ‘see-MOD-IPFV.HUM’ were rejected by my consultant. Modal passive forms usually denote epistemic modality (‘X can be done’). Examples (28) – (29) illustrate positive verb forms and examples (30) – (31) illustrate negative verb forms.

- (28) wàgàti^L gémbé kwè-lí dá:lá ʔwà=-mà-s-é=:
 sometimes other DEM-ACC outside see-MOD-FUT-3PL
 ‘Sometimes you can see them outside’.
- (29) kándá ʔyé: bálà-nàmà càm̀bà-mà-s-é=:ⁿ
 now today savannah-meat find-MOD-FUT-3PL=PTCP
 bè cè^L gèlè^L wá
 2PL POSS zone LOC
 ‘The wildfowl you can find now in our region...’
- (30) ʔyé: fwéré pyè-pyéw
 completely now completely
 k-wá: dèʔè-má-l-è ñ cùgó ìná
 DEM-LOC meet.↑-MOD-FUT.NEG-3PL1SG.EMPH DEM do.not.know
 ‘And so that you couldn’t find them... no, I don’t know such a thing’.
- (31) kò ^Hlwé ^{LHy}èlé: byà=-mà-ló
 NONH.POSS medicine easy have-MOD-FUT.NEG
 ‘The medicine for that cannot be found easily’.

The forms containing *-má-* are modal forms, in which the modal meaning cannot be expressed by the stem, which is simply an *a*-stem of verb. The Futurum/Habitual suffix, on the contrary, is able to express the semantics of the possibility, since it has a wide range of irrealis uses (see Dyachkov ms. for details). Thus, the modal value component can be introduced not by the *-má-* suffix, and we suggest that it is contributed by the aspect marker.

Syntactically, clauses containing modal passives are clauses lacking subjects. NPs which are adjacent to verbs are not likely to be subjects and do not trigger any verbal agreement. For example, in (32) the NP ‘medicines’ is plural but the verb is singular.

- (32) [kò ^Hlwé wé] ^{LH}yélé: **byà=-mà-ló**
 NONH.POSS medicine PL easy be.able-MOD-FUT.NEG
 ‘Medicines for that couldn’t be found easily’.

I consider the lack of overt subject to be the crucial syntactic property of *má*-forms. Some examples showing that sentences containing the overt subject are infelicitous are given below. Subject is present in (34) leading to an ungrammatical sentence.

- (33) séliyà:^L wá jò:ⁿ ló: jéʔⁿé byà=-mà-só
 custom LOC woman 2 marry.PFV be.able-MOD-FUT
 ‘According to our custom, one can marry two women’.
- (34) *séliyà wá ʔ jò:ⁿ ló: jéʔⁿé byà=-só
 custom LOC 1SG woman 2 marry.PFV be.able-FUT
 Int.: ‘According to our custom, I can marry two women’.

Semantically, the modality marker can express a range of modal meanings including deontic necessity, possibility etc. The following textual examples show that deontic modality can be expressed by the modal marker as well (35).

- (35) ká lá cé kòʔ-èⁿ kó kùgò cánd-á:
 QUOT LOGO POSS cauri-DIM INT hen middle-LOC
 là=-mà-s-é
 spend.night-MOD-FUT-3PL
 ‘He said: It is needed that my cauri spent the night among the chicken’.

At the same time, the participant-external possibility, which is naturally expressed by the verb *byàlà* in its Futurum/Habitual form, is not compatible with *má*-forms. However, this is due to the fact that modal marker is incompatible with overt subjects:

- (36) à dò:-(^{*}mà)-só ʔwà:nà dùmè
 2SG aller-MOD-FUT work finish.PFV
 ‘(Letting the addressee go.) You can go, the work is done’.

In a similar manner, the participant-internal possibility is also expressed by Futurum/Habitual but not *má*-forms:

- (37) *ɲ̀ là:rá lwí yà=:-(*)mà)-só
 1SG outside enter.PFV go-MOD-FUT
 Int.: ‘I may go outside’.

The semantics of epistemic modality is not compatible with *má*-forms. The most natural way to express the epistemic possibility or probability is the construction with the verb of knowing, in which the clause is marked by the question particle *má* (38). Modal passives cannot be used in such contexts (39).

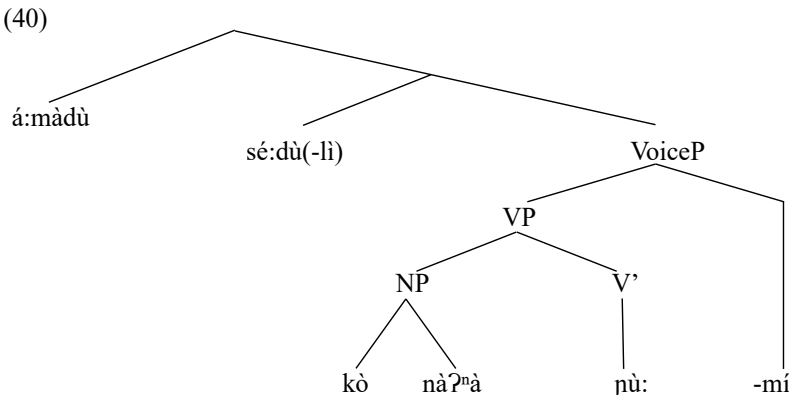
- (38) gèmbè cé ^{HL}sé:dù: mà iná
 other POSS S Q do.not.know
 ‘(Somebody is in the yard.) It may be Seydou’.

- (39) *sè:dú yàlà-mà-só
 S. come-MOD-FUT
 ‘(Somebody is in the yard.) It may be Seydou’.

Aforementioned types of modality are considered to constitute the core of the modality semantic field, (Plungian & van der Auwera 1998). However, I have shown that only a small set of meanings among these types are accessible to *má*-forms. It seems that the most appropriate contexts for modal passives are contexts in which epistemic or deontic modality (mostly participant-external) is expressed. Syntactically, modal passives are structures which lack true subjects since verb arguments do not agree in number with subjects. In the next section, I will try to show possible structural and semantic links between the modal passive and the causative.

5. Possible diachronic development

In section 3 I argued that causatives are monoclausal structures in which the presence of the internal agent (=causee) is obligatory and it cannot be co-indexed with the sentential subject. The correct constituent structure is represented below in (40).

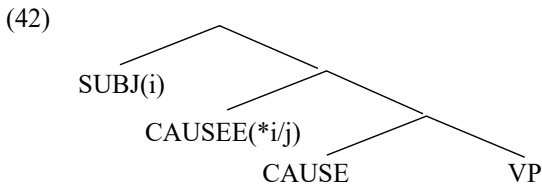


àmàdú sè:dù(-lí) kò ^Hnáʔná jù:-mí
 Amadou Seydou-ACC DEM cow eat-CAUS.D.PFV
 ‘Amadou made Seydou feed the cow’.

In this structure, the internal agent (=causee) is assigned the Accusative case by the Voice head (=the causative marker). However, in the discussion above, I have shown that the causee and the causing sub-event itself must be separated from each other. Strictly speaking, the causee does not have to be obligatorily used in such constructions, which is evidenced by examples like (23), repeated here as (41).

(41) sàtá há nàʔná jù:-mí
 S. REFL cow eat-CAUS.PFV
 ‘Sata_i fed her_i cow’.

The mechanism of how the causee is projected and receives case-marking needs additional investigation. However, I have shown that the causee is strongly associated semantically with the causative affix and is usually presupposed when the causative affix is used. Nevertheless, the sentential subject cannot be co-indexed with the causee, as I showed above. This fact is depicted in the following tree representing the abstract syntactic structure of the *má*-causative.



This observation is a strong argument against the scenario for the causative-to-passive shift of Haspelmath (1990), which is presented again in the examples (43)-(45), since the co-indexation of the subject and the cause is not possible.

- (43) *I have the barber shave me.* = ‘I CAUSED the barber to shave me’
 (44) *I have myself shaved by the barber.* = ‘I CAUSED MYSELF to be shaved by the barber’.
 (45) *I am shaved by the barber.*

A similar scenario was proposed by Say (2013) for the Kalmyk language. In this language causative constructions may have passive meanings as well:

(46) tuula čon-də id-ülə-v
 hare wolf-DAT eat-CAUS-PST
 ‘The hare was eaten by the wolf.’ (Say 2013: 276)

Say (2013) argues that “causative morphology does not necessarily signal the introduction of a separate causing subevent as such, but rather is used to shift the prominent

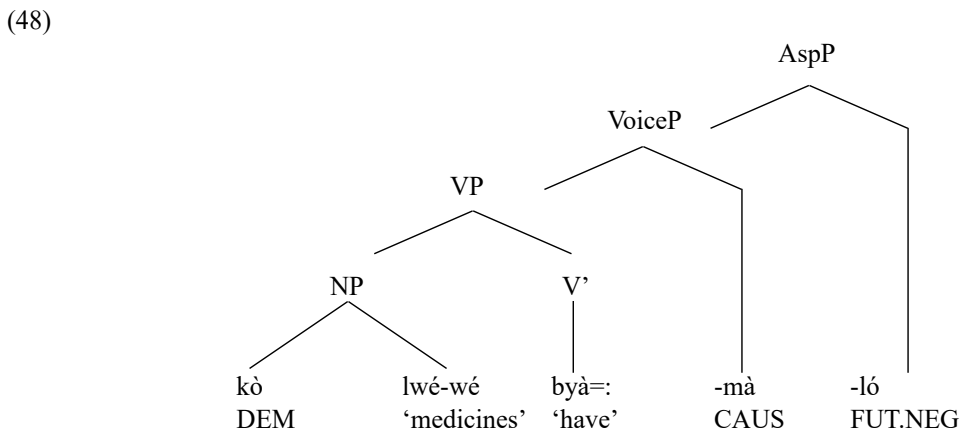
status of subject to a participant that is already present in the basic construction” (Say 2013: 277). Under this interpretation, clauses like (46) contain a covert argument which must be co-indexed with one of the participants present in the clausal structure (‘the hare made wolf eat X’). Naturally, this co-indexed argument X is pragmatically associated with the individual ‘hare’.

However, I suppose that these patterns cannot explain the Tomo Kan causative/passive polysemy and the main reason for this suggestion is that no co-indexation of arguments (let be they covert or overtly expressed) is possible. The exact mechanism of this phenomenon should be studied in a more thorough way. As for now, I claim that this cannot be due to structural properties of the causative. According to general principles, the causee would be able to be bound by the sentential subject and, in fact, it can do so if the possessor of the causee coincides with the subject, see (24). Thus, it is hardly plausible that the two arguments which are distinct would be reinterpreted as identical.

Another important feature of the passive marker is that the direct object is not structurally promoted, that is, it does not have a subject status. If this was so, number agreement would be triggered by the subject. However, we have seen that it is not the case since the plural marker does not require plural verb agreement.

- (47) kò ^Hlwé wé ^{LH}yèlé: **byà=-mà-ló**
 NONH.POSS medicine PL easy be.able-MOD-FUT.NEG
 ‘Medicines for that couldn’t be found easily’.

The most accurate syntactic representation for (47) would be the tree in (48).



In this structure, the direct object is not promoted to the position of the clausal subject. The crucial difference between the structure in (48) and the causative structure in (40) is that the former does not have any overt subject. However, we have seen before that it can be reconstructed as a generic entity (‘anyone’, ‘you’ etc.) which is able to perform the action denoted by the verb. Thus, a question arises whether this subject is present in syntactic structure, but I will leave this question open, since additional diagnostics is needed to detect its presence. However, the key difference between the causative and the passive is

the lack of surface subject, which facilitates the pragmatically-oriented reinterpretation of the structure. In addition, it is likely that the *-má* suffix cannot be adequately described as a marker deriving only distant causatives. In fact, only contact causatives result in a distant interpretation when combined with the *-má* suffix but other verbs do not specify the relation between the causing situation and the caused one. For example, it is unlikely that the verb ‘feed’, which comprises *-má* is necessarily a distant causative, since the event of feeding somebody can be performed both contactly and distantly. As a consequence, I put forward the hypothesis that the “distant” causative marker is not a marker which introduces the causing subevent but may be related to another grammatical notion, probably the voice head. This suggestion needs further investigation and elaboration.

6. Conclusion

I assume that the crucial difference between the causative and passive uses of the *má*-marker in Tomo Kan is the absence of the overt subject and the fact that the object is not promoted in passive structures. The main argument in favor of such analysis is the fact that pronouns which can bear accusative marking preserve it in modal passive constructions. Thus, the interconnection between the two uses can be described in terms of mechanisms involving subject deletion (ellipsis or pro-drop) rather than in terms of reflexivization or pragmatic reinterpretation, as it was proposed by Haspelmath (1990). I proposed that the modal component of the modal passive construction is introduced by the Future/Habitual marker, and, thus, the core structural difference between the causative and the passive is the lack of overt subject in the latter case. This feature is the main parameter differentiating the two uses. However, the difference is formulated in syntactic terms and leaves open the question of how the participant-external modality (expressed in passive clauses) is semantically associated with the meaning of distant causative, and this problem needs additional theoretical investigation.

Abbreviations

=	vowel lengthening	L	– low tone contour overlay
↑	– upstep	LOC	– locative
1,2,3	– 1 st , 2 nd , 3 rd person	LOGO	– logophoric pronoun
ACC	– accusative	MOD	– modal marker
CAUS	– causative	NEG	– negation
D	– distant	NONH	– nonhuman
DAT	– dative	PFV	– perfective
DEM	– demonstrative	PL	– plural
DIM	– diminutive	POSS	– possessor
EMPH	– emphatic form	PST	– past tense
FUT	– Futurum/Habitual	PTCP	– participle
HL	– high-low tone contour overlay	Q	– question particle
ID	– identification construction	QUOT	– quotative
INT	– intensifier	REFL	– reflexive
IPFV	– imperfective	SG	– singular

References

- Aissen, Judith. 2003. "Differential object marking: iconicity vs. economy". *Natural Language & Linguistic Theory* 21(3), 435-483.
- Comrie, Bernard. 1981. *Language universals and linguistic typology: Syntax and morphology*. Oxford: Blackwell.
- Dyachkov, Vadim (ms.). A grammar of Tomo Kan Dogon. Accessible at: https://www.academia.edu/39719090/A_Grammar_of_Tomo_Kan_Dogon
- Haspelmath, Martin. 1990. "The grammaticization of passive morphology". *Studies in Language* 14(1), 25-71.
- Kormushin, Igor. 1976. "O passivnom znachenii kauzativnykh glagolov" [Towards the passive meaning of causative verbs]. In: *Turcologica* Leningrad: Nauka, 89-93.
- Kulikov, Leonid. 2001. "Causatives". In: Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.). *Language Typology and Language Universals: An International Handbook*, Vol. 2. Berlin: Walter de Gruyter.
- Léger, Jean. 1971. *Grammaire dogon: Tomo-kan*. Bankass, Mali: Mission Catholique.
- Plungian, Vladimir & Johan van der Auwera. 1998. "Modality's semantic map". *Linguistic Typology* 2, 79-124.
- Rhee, Seongha & Hyun Jun Koo. 2014. "Grammaticalization of causatives and passives and their recent development into stance markers in Korean". *Poznań Studies in Contemporary Linguistics* 50 (3), 309-337.
- Say, Sergey. 2013. "Kalmyk causative constructions: case marking, syntactic relations and the speaker's perspective". *Suomalais-Ugrilaisen Seuran Aikakauskirja* [Journal de la Société Finno-Ougrienne] 94, 257-280.
- Yap, Foong Ha & Shoichi Iwasaki. 2003. "From causative to passive: A passage in some East and Southeast Asian languages". In: Eugene H. Casad & Gary B. Palmer (eds.). *Cognitive Linguistics and Non-Indo-European Languages* [Cognitive Linguistics Research 18], Berlin: Mouton de Gruyter, 2003, 419-446.

Asymmetries in negation in the Atlantic languages: a preliminary comparison

Abstract

This paper deals with negation in the Atlantic languages (Niger-Congo phylum), focusing on asymmetric negation that takes place when the differences between the negative and positive sentence go beyond the addition of the negative marker. The structural differences between positive and negative clauses show that negation cannot be analysed separately from the TAM system. Some observations on diachronic evolutions concerning TAM-negation interaction are also highlighted.

Keywords: Atlantic languages, negation, typology, TAM system.

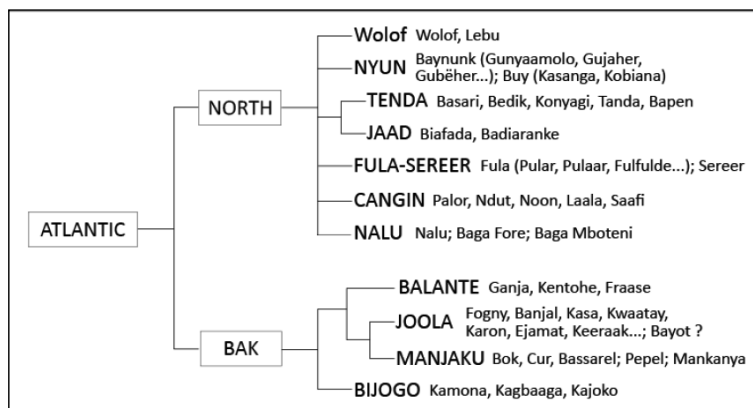
1. Introduction

Negation is one of the universal elements of human languages. Cross-linguistic studies have highlighted that negation is a complex process involving different grammatical categories, such as person and number marking, as well as the syntactic organisation of the sentence. Following Miestamo's (2005: 51) terminology, "when there are structural differences, i.e. asymmetry, between the affirmative and the negative in addition to the negative marker(s), the structures are asymmetric. Asymmetric negation can be divided into different subtypes according to what kind of asymmetries are found between the affirmative and the negative"¹. This paper provides a typological perspective, based on Miestamo's typology of negation (2005), focused on the types of asymmetric negation in the Atlantic languages. The Atlantic languages are part of the Niger-Congo phylum, which is subdivided into two main branches: the Northern and the Central or Bak ones².

The Atlantic languages are spoken in West Africa (mostly in Senegal, Gambia, Guinea and Guinea Bissau) and are known for their complex verbal morphology and heterogeneity. Due to the fact that these languages are unequally described, the analysis of negation was combined with a descriptive process, including fieldwork to complete the data. A description of the Tense-Aspect-Mood (TAM) system of the under-described languages was the first step leading to a typological study. The collected data comes from previously published studies and from the fieldwork.

¹ By opposition, a symmetrical negation is "when no structural differences are found between the affirmative and the negative in addition to the negative marker(s)" (Miestamo 2005: 51).

² Mel, Limba and Sua will be left aside; the focus is on "core" Atlantic.



Classification of the Atlantic languages (Pozdniakov & Segerer, to appear: 4)

This paper aims to explicit the organization of the TAM system, and sets an overview of the diversity of asymmetric negation in the Atlantic languages. In order to propose a cross-linguistic perspective, this paper will present a sample of 31 Atlantic languages, which are key languages in the different subgroups of the Atlantic, following Pozdniakov and Segerer's classification (to appear)³.

In the first part, we will discuss cases of constructional asymmetric negation, while the second will focus on paradigmatic asymmetric negation. Finally, we will show how other elements may be salient under negation.

2. Constructional asymmetric negation in the Atlantic languages

Constructional asymmetry takes place when a positive structure is different from its negative counterpart. Miestamo (2005) set up a precise typology of asymmetry, describing the Atlantic languages. It became obvious that several subtypes occur simultaneously, which also complicates the verbal morphology.

2.1. Some cases of TAM marking asymmetry

The TAM marking asymmetry can be divided into several subtypes. Miestamo (2005) defined different cases falling within the TAM marking asymmetry. This type of asymmetry is the most represented one in the Atlantic languages, so we will only present some characteristic examples. Concerning the TAM system in the Atlantic languages, the main aspectual opposition is between perfective and imperfective. The imperfective aspect may be divided into sub-aspects. It is relevant here to notice how one aspect is marked in positive clauses and what the negative counterpart of these clauses is.

³ List of languages can be found in the Appendix.

2.1.1. TAM marker replacement

The first type of asymmetry represented in the Atlantic languages is the replacement of the TAM marker by a ‘portmanteau’ marker: “the TAM marker is replaced by a marker which carries the meanings of the negative and the respective category TAM” (Miestamo: 2005, 116).

In Joola Banjal (Central Atlantic, Bak), the TAM marking asymmetry by replacement occurs to negate the future: the positive TAM marker (*pan*) is replaced by a negative TAM marker (*mat(i)*), as shown in example (1).

(1a)

pan
FUT

 i-tiŋ ‘I will eat’
1SG-eat

(1b)

mat(i)
NEG.FUT

 i-tiŋ ‘I won’t eat’
1SG-eat

(Bassène 2006; 2009)

In Joola Banjal, the habitual aspect shows how the nature of the TAM marker may differ in positive and negative sentences. In the positive ones, the habitual aspect is marked by the reduplication of the verb stem and the use of the epenthetic vowel (cf. 2a). In negative sentences the reduplication does not occur; instead there is a suffix carrying the habitual and the negative meanings (cf. 2b).

(2a) i-

tiŋ-c-tiŋ

 ‘I usually eat’
1SG-

eat-EPH-eat

(2b) i-tiŋ-

erit

 ‘I usually do not eat’
1SG-eat-

NEG.HAB

(Bassène 2006; 2009)

The replacement of the positive TAM marker by a ‘portmanteau’ marker is a very productive type of TAM marking asymmetry in the Atlantic languages. It may comprise affixes, independent morphemes or even the constructional marker. This high productivity highlights that in the Atlantic languages negation is integrated into the TAM system.

In Senegalese Pulaar, the subjunctive paradigm is marked by the morpheme *yo* (optative subjunctive) or *ma* (expectative subjunctive, cf. 3a); in negative sentences these markers are replaced by the prohibitive morpheme *hoto*, as shown in example (3b).

(3a)

ma
expt

 a loot ‘you have to wash’
2SG wash-ACT

(3b)

hoto
proh

 a loot ‘you don’t have to wash’
2sg wash-ACT

(Mohamadou 2014: 69)

2.1.2. Copula

Another well represented type of TAM marking asymmetry is the use of a completely different copula in positive and negative clauses (Miestamo 2005: 117). In the Atlantic languages, the negation process often replaces a positive copula with a negative one.

In Joola Banjal (Central Atlantic, Bak), the progressive aspect is marked by the use of a specific copula affixed by the subject pronoun, and associated with the infinitive form of the verb (cf. 4a). In the negative progressive, another copula occurs, replacing the positive one. The negative copula requires the change of the subject pronoun, and involves a different linear order of the constituents (cf. 4b and 4c).

- (4a) (ni)fi-tij

n-em
1SG-COP

 ‘I am eating’
INF-eat
- (4b)

i-let
1SG-COP.NEG

 nifi-tij ‘I am not eating’
1SG-COP.NEG INF-eat
- (4c)

i-lel-ene
1SG-COP.NEG-PAST

 nifi-tij ‘I was not eating’
1SG-COP.NEG-PAST INF-eat

(Bassène 2006; 2009)

In Buy Kobiaana (North Atlantic, Buy-Nyun), the imperfective clauses present an analytic structure using the copula *ɲgo* in the positive. In negative clauses, the copula occurring is *yul*, which I analyse diachronically as a result of the copula *ɲgo* with a negative suffix *-*VI* (the standard negation morpheme is *-il*), but it cannot be analysed synchronically. The grade of the verb stem’s initial consonant also differs under negation, this phenomenon will be discussed below (cf. section 3).

- (5a)

wo-ɲgo
3SG-COP

 rabb-a ‘he is writing’
3SG-COP I.write-IPFV
- (5b)

wo-yul
3SG-COP.NEG

 ddabb-a ‘he is not writing’
3SG-COP.NEG III.write-IPFV

(Montébran, fieldnotes 2017)

In Biafada (Nothern Atlantic), the progressive aspect is characterized by a gerund construction (marked by the suffix *-ä* attached to the verb). The subject is marked by a suffix placed after the gerund suffix, so that the structure of the clause is V-GER-S. The gerund form of the verb is thus marked for person. In the positive, the 3rd person plural pronoun presents the form *gəmmä*, which Wilson (1993: 72) analyses as a clue of an underlying structure based on the copula *ga* ‘to be’ and the pronoun *-mmä* 3rd person plural. The negative equivalent presents the negative morpheme *-al-* which is also used as a suffix for the negative perfective. The suffix appears independently of the verb. In the negative clause it is attached to the pronoun, although it is not a finite element. Wilson’s hypothesis is the following: the underlying form of the copula *ga* bears the negation marker *-al-* and

the person pronoun. Initial morpheme *g-* (cf. 6d) that would be remainder of the copula *ga* is not systematically attested. The structure is therefore type (COP)-NEG-S II.-GER (Wilson 1993: 72), as shown in example (6).

- (6a) reeg-ä-m ‘I am coming’
II.come-GER-1SG
- (6b) **al**-m reeg-ä ‘I am not coming’ or ‘I usually do not come’
NEG-1SG II.come-GER
(Wilson 1993: 72)
- (6c) reeg-ä-gəmmä ‘They are coming’
II.come-GER-3PL
(Wilson 1993: 72)
- (6d) (g)aal-e reeg-ä ‘(the man) is not coming’
COP-NEG-3SG II.come-GER or ‘(the man) usually do not come’
(Wilson 1993: 72)

This example of Biafada shows how the study of negation and its asymmetry gives an access to observing language change and to the diachronic evolution of the language.

2.1.3. A marker not used in the positive is added in the negative sentences

Another kind of asymmetry, much less common, is the addition of a morpheme not used in the positive clauses, which is not a negative morpheme (Miestamo 2005: 118). In Mankaña (Central Atlantic, Bak), the negative structure involves several elements: a specific subject pronoun (different that the subject pronoun in affirmative clause, cf. 7a and 7b), the addition of the negative morpheme (a prefix -VN- where N is a homorganic nasal), and the aspect marking change. In positive sentences, the perfective aspect is marked by the suffix *-i*, but the suffix is usually dropped in negative sentences. The imperfective aspect is not marked in simple positive clauses⁴ but is obligatory marked in negative clauses, as shown in examples (7).

- (7a) *d-lɛ:mp-i* ‘I worked’
1SG-work-PFV
- (7b) *m-ən-lɛ:mp-∅* takal ‘I did not work yesterday’
1SG.NEG-NEG-work-PFV yesterday
- (7c) *d-lɛ:mp-∅* ‘I am working’
1SG-work-IPFV

⁴The imperfective aspect is also marked in positive in some specific cases: such as in serial verb constructions.

- (7d) *m-əŋ-k̩-lɛ:mp* ‘I am not working’
1SG.NEG-NEG-IPFV-work
 (Montébran, fieldnotes 2016)

This type of asymmetry is really uncommon, so that it puts forward many hypotheses about TAM marking in negative sentences, which are reinforced by the example of Basari in the following section.

2.1.4 More complex TAM asymmetry

The Atlantic languages also present some cases, which are difficult to classify. In Basari (North Atlantic, Tenda), the aspectual opposition is marked by the place in the clause of a complex element, associating the paradigm marker for positive present indicative (*k'*) and the subject marker. If this element appears to the right of the verb, this is a perfective structure. If it appears to the left of the verb, this is an imperfective structure.

→ Perfective structure: V-[IND-S]

→ Imperfective structure: [IND-S]-V

In negative clauses, the paradigm marker *k'* does not appear, instead we see the negative discontinuous marker *a...-ná*. The first element of this negative morpheme is followed by the subject marker, then there is a verb, and eventually an aspect marker. Therefore, the linear order of the negative clause (NEG-S-V-TAM-NEG) completely differs from linear order of elements in the positive clause. Since the linear order of constituents does no longer distinguish the perfective and imperfective aspects, specific aspect markers occur: *-é-* to mark the perfective aspect (cf. 8b). The imperfective aspect is not marked, as shown in example (8d).

- (8a) *fuli-k̩-mé* ‘I talked’
talk-IND-1SG
- (8b) *a-xó-ʃamb-é-ná* *mɛlú-in* ‘He did not eat the rice’
NEG-3SG-eat-PFV.NEG-NEG *rice-DEF*
- (8c) *k̩-ó-yów* ‘He is coming’
IND-3SG-come
- (8d) *a-xó-yów-ná* *gər* *gáko* ‘He is not coming now’
NEG-3SG-come-NEG *LOC* *now*
 (Perrin 2019: 70-71)

Polarity in Basari reveals two radically different strategies of aspect marking. The opposition between positive and negative is highly differentiated from the structural point of view: the asymmetry concerns the whole information structure (negation appears as another mode, separate from the indicative).

2.2. Finiteness asymmetry

Finiteness asymmetry is characterized by the verb losing its finiteness. In most cases, a new finite element occurs in the clause – an auxiliary verb or a copula. As noticed by Dahl (1979), the negation tends to be attached to the finite element. The prototypical example of finiteness asymmetry is a positive synthetic construction and a periphrastic one as its negative equivalent.

The future in Laalaa (Lehar; North Atlantic, Cangin) is formed with the suffix *-an* attached to the verb in positive clauses, as shown in example (9).

- (9) mi ñam-an haay ‘I will eat couscous’
 1SG eat-FUT couscous

(Montébran, fieldnotes 2016)

Negative future clauses are built with the introduction of an auxiliary in negative form: *hay* (*hac+ri*), as shown in example (10).

- (10) mi hay ka-ñam jen ‘I won’t eat fish’
 1SG NEG.FUT INF-eat fish

(Montébran, fieldnotes 2016)

However the negative future can also be built by the standard negation suffix *-ri*, as shown in example (11).

- (11) mi ñamandi haay ‘I won’t eat couscous’
 mi ñam-an-ri haay
 1SG eat-FUT-NEG couscous

(Montébran, fieldnotes 2016)

This particularity should be looked upon with regard to the entire Cangin group. In five languages belonging to the group, the future can be expressed either by an analytic construction, or by a synthetic one, as shown in the following table:

	Laalaa	Noon	Ndut	Palor	Safen
Positive synthetic constructions	S V-FUT	S V-FUT (+certain)			
Positive Analytic constructions		S AUX PRON.O V.INF Aux: <i>hay</i> (-certain)	S AUX PRON.O V Aux: <i>ay</i>	S AUX PRON.O V Aux: SG <i>ʔaf</i> / PL <i>ʔayɔ</i>	S AUX V Aux: <i>hay</i>
Negative synthetic constructions	S V-FUT-NEG Standard Neg: <i>-ri</i>	S V-FUT-NEG Standard Neg <i>-ri</i> (+certain)			

Negative analytic constructions	S NEG.FUT INF.V Neg: <i>hay</i>	S AUX-NEG PRON.O V.INF Aux-Standard Neg: <i>hay-ki</i> (-certain)	S NEG.FUT V Neg : <i>yih</i>	S NEG (PRON.O) V Neg: <i>dii-jii</i>	S AUX-NEG V Aux-Standard Neg: <i>hay-dî</i>

Another example of finiteness asymmetry is provided by Balante Kantohe (Central Atlantic, Bak), where the negative marker is the morpheme *wot*. Synchronically, it seems to be difficult to define the nature of this morpheme. But its presence in negative sentences leads to a word order change; since the introduction of a finite element may attract other linguistic elements (i.e. object pronoun), we suggest the hypothesis that *wot* originally comes from a verbal stem, as shown in example (12).

- (12a) (ha)

<i>ñin</i>
look-PFV

ñi 'he looked at me'
(3SG) 1SG.O

(Doneux 1984: 66)

- (12b) (ha) **wot** *ñi*

<i>ñin</i>
look-PFV

 'he didn't look at me'
(3SG) **NEG** 1SG.O

(Doneux 1984: 66)

2.3. Person marking asymmetry

In the Atlantic languages, different types of person marking asymmetry occur. In the first type, negative clauses use a completely different set of markers (which occurs only in negatives), as shown in the example from Mankaña (7). The second type is the use of one set of markers selected among all sets and also used in positive sentences. This type occurs in Joola Banjal (Central Atlantic, Bak), where negation is marked by the replacement of the positive aspectual marker with negative “portemanteau” marker. It consists of a suffix *-ut* coding perfective aspect and negative meaning and the replacement of a subject pronoun, as shown in example (13). The set of subject pronouns used in the negative sentences also appears in other positive paradigms.

- (13a)

<i>ni</i>
eat-PFV

-tɨŋ-e 'I ate'
1SG

- (13b)

<i>i</i>
eat-NEG.PFV

-tɨŋ-ut 'I did not eat'
1SG

(Bassène 2006; 2009)

The person marking change is frequent in the Joola group⁵. It also seems to be common in the Bak group, as shown in the example of Mankaña (7).

⁵ Banjal, Keraak, Fogny, Gusiilay, Kasa.

3. Paradigmatic asymmetry

The paradigmatic asymmetry is defined as the non-one-to-one equivalence of positive and negative structures (Miestamo 2005: 54). In Biafada (North Atlantic), the imperfective is divided into subtypes: future, habitual and progressive. In positive clauses, the future and habitual meanings use the same structure whereas the progressive has its own construction. But in negative clauses, the habitual aspect falls within the progressive construction, whereas the future has its own construction.

	Future	Habitual	Progressive
Positive			
Negative			

(Following Wilson 1993: 72)

The positive paradigm (habitual / future) is marked by the auxiliary *nd* on which the negative suffix *-add-* may appear, but the negative paradigm only has a future meaning. To express the negative habitual, the speakers use the negative progressive, as shown in examples (14) and (15).

(14a) **nd-m** reegə ‘I will come’ or ‘I usually come’
 FUT-1SG II.come-VF
 (Wilson 1993: 72)

(14b) **nd-add-m** reegə ‘I won’t come’
 FUT-NEG.FUT-1SG II.come-VF
 (Wilson 1993: 72)

(15a) reeg-ä-m ‘I am coming’
 II.come-GER-1SG

(15b) **al-m** reeg-ä ‘I am not coming’ or ‘I usually do not come’
 NEG-1SG II.come-GER
 (Wilson 1993: 72)

This kind of paradigmatic asymmetry leads to reattribution of semantic value. In other languages, like in Senegalese Pulaar, the paradigmatic asymmetry leads to the neutralisation of positive distinctions. Indeed, the positive perfective presents five different paradigms, which are all negated by only two negative constructions (three including the one to negate emphasis); the same holds for the imperfective, where five positive paradigms correspond to two negative constructions (three including the one to negate emphasis).

Such global paradigmatic asymmetry is not frequent in the Atlantic languages, the most cases of this kind of asymmetry concern only one construction.

4. Other things may be salient under negation

Negation is a complex process involving many linguistic elements. In some languages, the marked categories may be different in positive to negative clauses, even if they do not directly carry a negative meaning. For example, in Buy Kobiana (North Atlantic, Buy-Nyun) the consonant gradation has different roles in positive and in negative clauses.

The perfective aspect is marked by a synthetic verbal construction, ending in the suffix *-i*. The verb stem begins with a grade III consonant (which distinguish the perfective aspect from the imperfective one being characterized by the grade I consonant). The negative equivalent is also a synthetic verbal construction; the negative suffix *-il* is added after the perfective marker. Concerning the initial consonant of the verb stem, the 1st and 2nd person singular present a grade I consonant, while the rest of the paradigm presents a grade III consonant as shown below.

Positive Perfective			Negative perfective		
S-III.V-PFV			S-GRD.V-PFV-NEG		
1SG	ma- dd abb-i	‘I wrote’	ma-rabb-i- il		‘I did not write’
2SG	a- dd abb-i	‘you wrote’	a-rabb-i- il		‘you did not write’
3SG	wo- dd abb-i	‘he wrote’	wo- dd abb-i- il		‘he did not write’
1PL	ngee- dd abb-i	‘we wrote’	ngee- dd abb-i- il		‘we did not write’
2PL	kaa- dd abb-i	‘you wrote’	kaa- dd abb-i- il		‘you did not write’
3PL	naan- dd abb-i	‘they wrote’	naan- dd abb-i- il		‘they did not write’

(Montébran fieldnotes, 2017)

In the imperfective aspect, the verb shows a grade I consonant and ends with the suffix *-a*. The progressive sub-type is marked by a periphrastic structure containing the copula *ngo*. The negative equivalent presents the negative copula *yul* (historically probably *ngo-*yl* with consonant gradation). Concerning the initial consonant of the main verb stem, the 1st and 2nd person singular present a grade I consonant, while the other person present a grade III consonant.

Positive Imperfective			Negative Imperfective		
S-III.COP I.V-IPFV			S-I.COP.NEG GRD.V-IPFV		
1SG	ma-ngo rabb-a	‘I am writing’	ma- yul rabb-a		‘I am not writing’
2SG	a-ngo rabb-a	‘you are writing’	a- yul rabb-a		‘you are not writing’
3SG	wo-ngo rabb-a	‘he is writing’	wo- yul ddabb-a		‘he is not writing’
1PL	ngee-ngo rabb-a	‘we are writing’	ngee- yul ddabb-a		‘we are not writing’
2PL	kaa-ngo rabb-a	‘you are writing’	kaa- yul ddabb-a		‘you are not writing’
3PL	naan-ngo rabb-a	‘they are writing’	naan- yul ddabb-a		‘they are not writing’

(Montébran, fieldnotes 2017)

The summary of the grade of the stem consonant is presented below:

	Perfective		Imperfective	
	POS	NEG	POS	NEG
1SG	III	I	I	I
2SG	III	I	I	I
3SG	III	III	I	III
1PL	III	III	I	III
2PL	III	III	I	III
3PL	III	III	I	III

In the positive clauses, the perfective is marked by a grade III consonant, while the imperfective is marked by a grade I consonant for all persons. In the negative clauses, the 1st and the 2nd person in singular, i.e. the speakers present the grade I consonant, while the other persons (3rd person singular and all persons in plural) have a grade III consonant. This specific consonant gradation in the negative context leads to the marking of another linguistic elements, which distinguish the speakers from the other persons.

A similar case appears in Gujaaher (North Atlantic, Buy-Nyun), where in future tense the negative morpheme (and its position) differs between 1st/2nd person and the 3rd person (singular and plural). For the 1st and 2nd person, the negation is marked by the affix *-ga-* occurring between subject and verb; for the 3rd person, the negative morpheme is the prefix *b-* followed by the subject and the verb, as shown in the following table.

	Positive (S-V-FUT-(S))		Negative (S-NEG.FUT-V-(S) or NEG.FUT-3SG-V)	
1SG	man-ceem-o	‘I will sleep’	man- ga -ceeme	‘I will not sleep’
2SG	fu-ceem-o	‘You will sleep’	fu- ga -ceeme	‘You will not sleep’
3SG	a-ceem-o	‘He will sleep’	b -a-ceeme	‘He will not sleep’
1PL.INCL	ij-ceem-o-minij	‘We will sleep’	ij- ga -ceeme-minij	‘We will not sleep’
1PL.EXCL	ij-ceem-o-ndoj	‘We will sleep’	ij- ga -ceeme-ndoj	‘We will not sleep’
2PL	kan-ceem-o-kuj	‘You will sleep’	kan- ga -ceeme-kuj	‘You will not sleep’
3PL	an-ceem-o	‘They will sleep’	b -an-ceeme	‘They will not sleep’

(Montébran, fieldnotes 2017)

The above table shows that in Gujaaher the 1st/2nd person in singular and plural stands in opposition to the 3rd person (singular and plural).

5. Conclusion

The study of negative asymmetries points out the way negation interacts with other linguistic elements. It may concern the sentence construction (TAM marking, person marking, finiteness, etc.) or the whole verbal paradigm (no one-to-one equivalence); it may also feature other elements (i.e pragmatic categories).

This overview highlights the heterogeneity of the negative strategies and their interactions in the Atlantic languages, where negation is overtly marked and integrated into the TAM system. Within the Atlantic languages, the study of negation shows that some subgroups tend to develop similar negative strategies or interactions, which may also give some clues for the internal clustering in the Atlantic group.

Abbreviations

1	first person	IPFV	imperfective
2	second person	LOC	locative
3	third person	NEG	negation
ACT	active radical	O	object
AUX	auxiliary	OBL	obligative
COP	copula	OPT	optative
DEF	definite article	PAST	past
DEM	demonstrative	POS	positive
EPH:	epenthetic vowel	PFV	perfective
EXPT	expectative	PL	plural
FOC	focalisation	PREP	preposition
FUT	future	PROH	prohibitive
HAB	habitual	PRON	pronoun
GER	gerund	S	subject
GRD	consonant gradation	SG	singular
INF	infinitive	TAM	Tense-Aspect-Mood
IND	indicative marker	V	verb

Appendix. Languages sample

The languages used in this study are the following:

Biafada	Joola Karon
Badiaranke	Joola Kasa
Balante ganja	Joola Kerak
Balante Kantohe	Koñagi
Bassari	Laalaa (Lehar)
Bedik	Mankaña
Bijogo	Manjaku
Buy Kobiana	Ndut
Gujaahar	Noon
Gubeeher	Palor
Guñaamolo	Pepel
Joola de Carabane	Pulaar
Joola Banjal	Saafi-saafi
Joola Bayot	Sereer
Joola Fogny	Wolof
Joola Gusiilay	

References

- Bassène, Alain-Christian. 2006. Description du Joola Banjál (Sénégal). Ph. Thesis, Université Lumière Lyon 2, Lyon.
- Bassène, Alain-Christian. 2009. “La négation en Jóola Banjál”. *Sciences et Techniques du Langage* 6. CLAD: Dakar.
- Dahl, Osten. 1979. “Typology of sentence negation”. *Linguistics* 17, 79-106.
- Doneux, Jean Léonce. 1984. *Éléments de grammaire Balante, établis pour le département d’alphabétisation des adultes*. Bissau: Ministère de l’éducation nationale.
- Miestamo, Matti. 2005. *Standard Negation: The Negation of Declarative Verbal Main Clauses in a typological Perspective*. Berlin/NewYork: Mouton de Gruyter.
- Mohamadou, Aliou. 2014. *Le verbe en peul* (2nd ed). Paris: Karthala.
- Perrin, Loïc-Michel. 2019. *Description grammaticale du basari (oniyan): langue atlantique du Sénégal oriental*. Paris : Inalco presses.
- Pozdniakov, Konstantin & Guillaume Segerer. (to appear). “A Genealogical classification of Atlantic languages”. Luepke (ed.). *The Oxford guide to the Atlantic languages of West Africa*, Oxford: Oxford University Press.
- Wilson, William André Auquier. 1961. “Outline of the Balanta language”. *African Studies* 2, 139-168.

Let's tidy up the grammar of Hausa and analyse *after/behind* in *after class/after he arrives/he's behind* etc. as prepositions

Abstract

Hausa is generally analysed as having three distinct classes of preposition: (1) core monomorphemic, e.g., *à* 'in, at, on', *dà* 'with'; (2) "genitive" prepositions, e.g. *baaya-n* 'behind, after' (lit. 'behind/back of', cf. *baayaa* 'back'); (3) two-word transitive prepositions with *dà* 'with' as the second element, e.g. *kusa* 'near(by)' → *kusa dà* 'near to'. Under conventional classifications, prepositions such as the above are basically defined as governing NPs, e.g., *baayan kwaanaa ukù* 'after three days'. Conjunctions, however, traditionally govern subordinate clauses, e.g., *baayan sun isoo* 'after they arrived', and adverbs usually modify verbs, e.g., *sun koomàa baaya* 'they went behind'. Following early work by Jespersen (1909-49), extended more recently in Pullum & Huddleston (2002), this paper provides further evidence for a unified analysis which treats all the above as prepositional heads of phrases which, like verbs, simply differ in the complements they are licensed to take – NPs, clauses, or zero.

Keywords: intransitive prepositions, transitive prepositions, complementation

1. Introduction

The conceptual framework underpinning this paper departs from the traditional viewpoint and treats a number of prepositions as heads of phrases with different complements, structurally equivalent to phrases headed by NPs and verbs. Much recent work in descriptive grammar along these lines expands the prepositional category by re-assigning to it many traditional subordinating conjunctions and adverbs (see, for example, Emonds (1972), Jackendoff (1973), Pullum & Huddleston (2002), Pullum (2009), following Jespersen (1909-49: part 2: 10-15; 1924: 87-89). This paper examines the corresponding constructions in Hausa (Chadic, Afroasiatic), applying the same analytical framework. Prepositions in Hausa represent a closed class of relators of three main formal types, all sharply distinct: (a) nonderived monomorphemic prepositions, e.g., *à* 'in, on, at, etc.'; (b) so-called "genitive prepositions" consisting of a body-part noun with a genitive 'of' suffix, e.g., *baaya-n* 'behind, after' (lit. 'behind/back of', cf. *baayaa* 'back'); (c) two-word transitive prepositions formed with *dà* 'with' as the second element, e.g., *kusa* 'near(by)' → *kusa dà* 'near to'.¹

¹ My thanks to Hamisu Salihu for sharing his insights with me, and to Nina Pawlak for comments on an earlier version. Transcription and abbreviations: *à(a)* = low tone, *â(a)* = falling tone (on heavy CVV or CVC syllable), high tone is unmarked; *aa, ii*, etc. = long, *a, i*, etc. = short; *ɓ, ɗ* = laryngeal implosives,

Sections 2 and 3 survey conventional classifications and the central properties of Hausa prepositions, with some observations on how body parts such as ‘back, mouth, forehead, front of body, head’ are all pressed into service as orientational spatial prepositions. Section 4 looks at how Hausa scholars have classified prepositions. Finally, in Section 5 we present a range of data that show why many traditional subordinating conjunctions are best analyzable as prepositional heads of prepositional phrases. The supportive evidence lies in parallel constructions and complementation patterns associated with the three (a, b, c) types, i.e., governing either NP, clausal or zero complements.

2. Traditional classifications (English)

Before examining those items now subsumed under the label “preposition”, we first of all provide an overview of how English grammarians have described prepositions. In general, traditionally (and cross-linguistically) prepositions govern NPs, conjunctions govern subordinate clauses, and adverbs usually modify verbs. Distinct part-of-speech labels would therefore be assigned to each of the underlined words in (1a-c), i.e., prepositional ‘after three days’, subordinating conjunction ‘after you have left’, and adverbial ‘they went behind, at the back’. Thus:

1. a. If [+NP] as in ‘after [three days]’ then ‘after’ = *preposition*, but:
 - b. If [+clause] as in ‘after [you have left]’ then ‘after’ = *subordinating conjunction*
 - c. If [+ Ø] as in ‘they went [behind]’ then ‘behind’ = *adverb*

This paper broadly follows the detailed reformulation (of English) presented in Pullum & Huddleston (2002) and departs from the orthodox classification *by treating prepositions as structurally equivalent to phrases headed by NPs and verbs*, so analyzing *all* occurrences of ‘after’ and ‘behind’ in (1) as prepositional. Independent support comes from parallel syntactic configurations elsewhere in the grammar.²

3. The central properties of Hausa prepositions

As outlined above, Hausa has basically three formally distinct subtypes of preposition: (a) core, and derived (b) genitival and (c) two-word forms.

(a) Core prepositions

With the exception of **gà** ‘(in relation) to’ (= **gàree** pre-pronoun), core prepositions are all non-inflecting particles and are generally acknowledged to include: **à** ‘in, on, at,

k, **ts** = ejectives, **ʔy** = glottalized palatal glide, **ṛ** = apical tap/roll, **c** and **j** = palato-alveolar affricates. Abbreviations: COMP = complement; F = feminine; FOC.PFV = focus (narrative) perfective; FUT = future; IPFV = imperfective; IO = indirect object; M = masculine; PFV = perfective; PL = plural; REL = relative marker; SG = singular; SBJV = subjunctive.

² Bagari (1987), a Hausa-speaking linguist, extends the prepositional classification to clausal complements (probably reflecting his native-speaker intuitions), but he fails to explain how/why this analysis is at variance with the orthodox preposition-conjunction dichotomy.

etc.’, **dà** ‘with’, **dàgà** ‘from’, **ta** ‘via’, and **yà** (= **i**) ‘like’. They differ sharply from other prepositions. A key syntactic feature which distinguishes them from genitive prepositions is that they do not permit stranding, a restriction they share with two-word prepositions (see below). Nor do they freely govern content clauses (ex. 4 with **dàgà** ‘from’ is exceptional). Example (2) illustrates this subtype with the high-frequency general preposition **à** ‘in, on, at’.

2. *naa gan shì à kàasurwaa*
 1SG.PFV see 3MSG in market
 ‘I saw him **in/at** the market’

(2) has a preposition **à** in prototypical usage—with an object NP ‘market’ as complement. Sentences (3) and (4) provide an example of the unifying generalization. The source locative **dàgà** straightforwardly governs NPs as prepositional ‘from’:

3. *yaa zoo [dàgà Kanò]*
 3MSG.PFV come from Kano
 ‘he came **[from Kano]**’

Dàgà can also take clausal ‘just because’ complements, e.g.:

4. *[dàgà naa cée ’yàayà dai?’] sai ya taasoo mini*
 from 1SG.PFV say how then then 3MSG.FOC.PFV attack 1SG.IO
 ‘**[just because I said ‘how’s it going then?’]** he attacked me’

Within the revised framework both constructions contain the same prepositional head **dàgà** with different dependents.

(b) “Genitive prepositions”

The second subset contains so-called “genitive prepositions” (see especially Newman 2000: 466ff.; Jaggar 2001: 667ff.; also Kraft 1963: vol. 2; Pawlak 2007:101ff.; Wolff 1993: 439ff.). These are composed of a body-part noun with a genitive ‘of’ suffix, usually masculine **-n**. All can routinely take an NP object, and some have multiple complementation, presenting different complement patterns, not all of which have been traditionally described (see below).

A natural question to ask at this point is: what (if any) are the correlations between the anatomical location of the body-part term itself and the derived prepositional meaning(s)? On the back-interior axis, ‘back’ and ‘stomach’ are both sources for the orientational genitive prepositions **baayan** and **cikin**. **Baaya-n** (lit. ‘behind/back-of’, cf. **baayaa** ‘back (of body)’) expresses locative ‘behind/at back of’, and can take a prototypical NP complement:

5. *yanàa [baayan kujèeraa]* [= object NP COMP]
 3MSG.IPFV back.of chair
 ‘it’s **at the back of** the chair’

Prepositions can take PP complements. So in example (6) we have a head core preposition **dàgà** ‘from’ governing a PP **kân dookii** ‘on the horse’:

6. *taa faadî [dàgà kân] dookii*
 3FSG.PFV fall from on.of horse
 ‘she fell **from (on)** the horse’

Baayan can also express temporal ‘after’, with a complement perfective TAM clause (7):

7. *zân zoo [baayan sun taashi] [= perfective clause COMP]*
 FUT.1SG come after.of 3PL.PFV leave
 ‘I’ll come [**after** they have left]’

(With (7) **baayan** conventionally labeled as a conjunction in this environment.)

The genitive preposition **ciki-n** ‘in(side)’ encodes interior location/containment, and is related to the noun **cikii** ‘interior, inside’, and ultimately the body-part **cikii** ‘stomach’, e.g.:

8. *kaayaa sunàa cikin àkwààti*
 stuff 3PL.IPFV inside.of box
 ‘the stuff is **in(side)** the box’

Cikin is also extended into the non-spatial domain to denote an ongoing, unbounded activity, usually with an actional/eventive noun or verbal noun of some kind, e.g.:

9. *sunàa cikin tàfiyàa/kookawàa*
 3PL.IPFV inside.of travelling/wrestling
 ‘they were travelling along/wrestling’ (lit. ‘**inside of** travelling/wrestling’)

On the up-down axis, front body parts such as ‘mouth, forehead, front of body, head’ are all pressed into service as orientational prepositions. A number of these genitive prepositions subcategorize for both NP and clausal complements, licensing a synonymous alternation between a TAM-less verbal noun of some kind and a subjunctive TAM complement. The matrix clause typically occurs with a stative imperfective TAM. Examples (10-15) illustrate:

10. *yanàa bàaki-n [zuwàa/isôowaa]* [verbal noun]
 3MSG.IPFV mouth-of coming/arriving
 =
11. *yanàa bàaki-n [yà zoo/isoo]* [subjunctive]
 3MSG.IPFV mouth-of 3MSG.SBJV come/arrive
 ‘he’s **just about** to arrive’
12. *sunàa bàaki-n [shìgaa masallaacii], sai sukà hangoo sarkii*
 3PL.IPFV mouth.of entering mosque then 3PL.FOC.PFV see emir
 [verbal noun]

- =
13. *sunàa* *bàaki-n* [*sù shìga masallaacii*], *sai sukà hangoo sarkii*
 3PL.IPFV mouth-of [3PL.SBJV enter mosque] then 3PL.FOC.PFV see emir
 [subjunctive]
14. *inàa* *gàbàani-n* [*tàfiyàa*] *sai...* [verbal noun]
 1SG.IPFV front-of going then
- =
15. *inàa* *gàbàani-n* [*in tàfi*] *sai...* [subjunctive]
 1SG.IPFV front-of 1SG.SBJV go then
 ‘I was **on the point of** leaving when...’

Notice the strong [front = future] correlation between the front-of-body-derived temporal prepositions and the immediate futurity notions ‘on the point of, about to, just before’.

Non-body part terms with **dà** can also participate in the alternation, e.g.,

16. *tanàa* *dab dà* [*kammàlà aikìn*] [bare verb]
 3FSG.IPFV close to finish work.the
- =
17. *tanàa* *dab dà* [*tà kammàlà aikìn*] [subjunctive]
 3FSG.IPFV close to 3FSG.SBJV finish work.the
 ‘she’s **close to** finishing the work’

In contrast to core prepositions, the complement of genitive prepositions can be left unexpressed but recoverable from the context. If no complement is expressed, the preposition drops the /-n/ suffix and appears with a final short vowel, which contrasts with the cognate (body-part) nouns with a long final vowel, i.e., **baaya** ‘behind, at the back’ < **baayaa** ‘back of body’, **ciki** ‘inside’ < **cikii** ‘interior’, etc. (18) illustrates:

18. *sun* *koomàa* [*baaya*]
 3PL.PFV go to at the back/behind
 ‘they went [**to the back/behind**]’

The related short final vowel variant **baaya** ‘to/at the back, behind’ occurring with the verb ‘go to, return to’ in (18) has no NP complement and as such is traditionally classified as an adverb in both English and Hausa.³

³ Historical note: most common nouns now end in a long vowel. Historically, however, common nouns ended in short final vowels, but underwent lengthening, a change which did not affect traditional ‘adverbs’. Thus, although the noun **baayaa** ‘back’ and adverb **baaya** ‘at the back’ differ synchronically, the final short vowel on **baaya** is in fact a retention, so that the occurring body-part noun **baayaa** derives from the historically original adverbial form ***baaya**.

(c) Two-word transitive prepositions formed with **dà** ‘with’ as the second element

As noted in Jaggard (2010), some intransitive prepositions, e.g., **kusa** ‘near’, **neesà** ‘far away’, when used in transitive prepositional constructions take a complement headed by the fixed transitive preposition **dà** ‘with’ (functionally comparable to grammaticized genitive prepositions with ‘of’). (19) illustrates with **kusa dà/neesà dà X** ‘near to/far from X’:

19. *gidaanaa nàa [kusa dà/neesà dà masallaacii]*
 house.of.1SG IPFV near/far with mosque
 ‘my house is [**near to/far from** the mosque]’

Members of a small subset of (deverbal) stative formations also take the same **dà**-headed complement when functioning as transitive prepositions, e.g. **tàare** ‘together’ → **tàare dà** ‘together with’ (cf. **taaràa** ‘join, connect’). (20) illustrates:

20. *naa zoo [tàare dà yaarònaa]*
 1SG.PFV come together with son.of.1SG
 ‘I came [**(together) with** my son]’

Tàare ‘together’ is unusual in that it can occur without a complement, e.g.,

21. *mun zoo tàare*
 1PL.PFV come together
 ‘we have come **together**’

Again, because traditional prepositions require complements, **tàare** in (21) would be defined as an adverb.

Other common examples (all transitive prepositions) are: **dàngàne dà** ‘related to, regarding’ (cf. **dangànaa** ‘prop/lean against’), **fiye dà** ‘more than’ (cf. **fi** ‘exceed’), **gàme dà** ‘concerning, regarding’ (cf. **gamàa** ‘join, connect’), **hàde dà** ‘combined with’ (cf. **hadàa** ‘combine, join’).

Table 1 summarizes the analysis so far:

Feature	Core prepositions	Genitive prepositions	Two-word prepositions
Stranding	-	+	-
Monomorphemic	+	-	-

Table 1. Morphology and syntactic distribution of Hausa prepositions.

4. Recent descriptions in Hausa dictionaries, wordlists and grammars

Recent reference grammars and dictionaries basically repeat the same classification as outlined for English (see Caron 1987: 58ff., 140ff.; Wolff 1993: 439ff.; Newman 2000:

chap. 57; Jaggar 2001: 606ff.). Examples (22, 23) provide further illustration, correctly analyzing **kààfin** ‘before’ in (22) as a preposition:⁴

22. *yaa isoo [kààfin tsakar̃ raanaa]*
 3MSG.PFV arrive before middle.of day
 ‘he arrived [**before** midday]’

When followed by a complement consisting of a *declarative content clause*, however, the same lexeme **kààfin** would switch word class and be classified as a *conjunction* initiating a subordinate clause, e.g.,

23. *yaa isoo [kààfin raanaa tà faadi]*
 3MSG.PFV arrive before sun 3FSG.SBJV fall
 ‘he arrived [**before** the sun set]’

The data below are instructive and illustrate the occasional conceptual and classificatory uncertainties across a selection of recent Hausa-English dictionaries and grammars, showing how they can at times vary in their definitions. The non-exhaustive list includes some high-frequency elements with their basic meanings, and the prepositional (prep.) and/or conjunction (conj.) labels assigned to them.

(a) Core prepositions (usually only with NP complements)

à ‘in, on, at, etc.’ = prep: P. Newman & R. Newman (1977); P. Newman (2007); Wolff (1993).

dà ‘with’; ‘that’ (Jaggar (2010) ‘that’ = ‘subordinator’)

= prep. & conj. : P. Newman & R. Newman (1977); Wolff (1993)

= prep.: P. Newman (2007 = ‘particle’ with clausal objects)

dàgà ‘from’ = prep.: P. Newman & R. Newman (1977); P. Newman (2007); Wolff (1993)

gà ‘(in relation) to’ = prep.: P. Newman & R. Newman (1977); P. Newman (2007); Wolff (1993)

ta ‘via’ = prep.: P. Newman & R. Newman (1977); P. Newman (2007); Wolff (1993)

yà (= i) ‘like’ = prep.: P. Newman & R. Newman (1977); P. Newman (2007); Wolff (1993)

(b) Genitive prepositions (with varied complementation)

bàakin ‘on the point of, about to’ (cf. **bàakii** ‘mouth’)

= prep: P. Newman (2007, includes clausal COMP examples);

P. Newman & R. Newman (1977, includes clausal COMP examples); Wolff (1993)

baayan ‘behind, after’ (cf. **baayaa** ‘back of body’)

= prep. & conj.: Jaggar (1992); P. Newman (2007); Wolff (1993)

= prep: P. Newman & R. Newman (1977, with NP and clausal COMP examples)

⁴ Recent descriptions in other Chadic languages use essentially the same [preposition + NP] vs. [conjunction + clause] distinction, e.g., (Jungraitmayr 1991, Tangale); Leger 1994, Kwami); Schuh 1998, Miya); (Haruna 2003, Guruntum). In his Ngizim dictionary (1981) Schuh hints at something close to a unified prepositional analysis. Thus, he has “preposition” as an independent grammatical category, but no part-of-speech “conjunction”. Relevant dictionary entries are then described as governing a full sentence, an NP or a PP, usually without a category label.

cikin ‘in(side)’, ‘among’, ‘in the midst of’ (cf. **cikii** ‘stomach’)

= prep.: P. Newman & R. Newman (1977); P. Newman (2007, includes clausal COMP example); Wolff (1993)

gàban ‘in front of’ (cf. **gàbaa** ‘front of body’)

= prep.: P. Newman & R. Newman (1977); Jaggat (1992); Wolff (1993)

gàbànin/gàbànnin ‘on the point of’, ‘just before’ (see **gàban**)

gòoshin ‘just before’ (cf. **gòoshii** ‘forehead’)

= prep.: P. Newman & R. Newman (1977); Wolff (1993)

kân ‘on (top of)’, ‘regarding’ (cf. **kâi** ‘head’)

= prep.: P. Newman & R. Newman (1977); Jaggat (1992); P. Newman (2007); Wolff (1993)

(c) Two-word prepositions with **dà**

tàare dà ‘together with’

= adverb: P. Newman (2007); P. Newman & R. Newman (1977)

dab dà ‘very close to’

= adverb: P. Newman & R. Newman (1977); P. Newman (2007)

= conj.: Wolff (1993)

Recent reductionist approaches would achieve an important generalization by plausibly analyzing all instances of, for example, ‘after/at the back/behind’ and ‘before’ etc., as PREPOSITIONS which (like verbs) simply differ in the complements they are licensed to take.

5. Prepositions, as heads of phrases, govern different types of dependents (not just NPs)

We now propose a number of arguments which, when taken together, strongly favour the re-analysis.

5.1. Finite TAM clause = nominalized VP with same head preposition

The most decisive support for the re-alignment comes from alternations in complement type, specifically where the same prepositions allow a clause with a TAM to alternate with a reduced nominalized VP as a synonymous variant, e.g.,

24. a. *zân* *zoo* [*baayan sun taashi*] [= clausal COMP]
 FUT.1SG come after.of 3PL.PFV leave

=

b. *zân* *zoo* [*baayan taashinsù*] [= nominalized VP COMP]
 FUT.1SG come after.of leaving.of.3PL

‘I’ll come [**after** they have left]’ (lit. ‘...**after** their leaving’)

25. a. [*gàrin* *in* *rabàa su*] *sai na* *yi ràunii*
 in process of 1SG.SBJV separate 3PL then 1SG.FOC.PFV do injury
 [= clausal COMP]

28. a. *naa nɛɛmi [ɪn gan shi]*
 1SG.PFV seek 1SG.SBJV see 3MSG
 ‘I tried (sought) [to see him]’

Cf. the identical verb form with a direct object noun:

- b. *naa nɛɛmi [aikii]*
 1SG.PFV seek work
 ‘I looked for [work]’

Again we see a syntactic equivalence with prepositions — the same form of the verb taking a clausal (28a) and NP (28b) object. There is no justification for assigning the two instances of *nɛɛmi* ‘seek’ to different primary part-of-speech categories. Prepositions are best categorized in the same way.

5.4. WANT verb complements and NPs with genitive prepositions

Matrix WANT verbs in the imperfective can suffix the *-n* ‘of’ genitive clitic and so govern a genitive complement lower clause, e.g.,

29. *kinàa sô-n [ki tàfi kàasuwaa]?* [subordinate SBJV clause]
 2FSG.IPFV wanting-of 2FSG.SBJV go market
 ‘do you want [to go to the market]?’
 (lit. ‘you are wanting of you should go...?’)

This is the same bound *-n* ‘of’ suffix which genitive prepositions require when acting as transitive prepositions with NP complements, e.g.,

30. a. *naa gan shi [kàrkashi-n teebùr]*
 1SG.PFV see 3MS underneath-of table
 ‘I saw it [underneath (of) the table]’
 b. *yaa sâa jàkaa [baaya-n kujèeraa]*
 3MSG.PFV put bag back-of chair
 ‘he put the bag [behind the chair]’

Hausa marks both NP and clause complement-types with the same (genitive) linker.

5.5. Genitive prepositions can take zero complements (‘he’s inside’ = prep. + Ø)

The complement *ciki* ‘inside’ in an utterance such as *yaa shiga [ciki Ø]* ‘he went [inside]’ is routinely classified as an “adverb” in both Hausa and English, because traditional prepositions must have overt complements. Examples (31-34) further illustrate:

31. *sun koomàa [baayan gàrii]*
 3PL.PFV go to back/behind.of town
 ‘they went [outside the town]’ (lit. ‘**back/behind of** town)

Cf:

32. *sun koomàa [baaya Ø]*
 3PL.PFV go to back/behind
 ‘they went [to the back/behind Ø]’

33. *yanàa [cikin ajii]*
 3MSG.IPFV inside.of classroom
 ‘he’s [in(side) the classroom]’

Cf.

34. *yanàa [ciki Ø]*
 3MSG.IPFV inside
 ‘he’s [in(side) Ø]’

In (32) and (34) the related (short final vowel) forms **baaya** ‘to the back, behind’ and **ciki** ‘inside’ have no overt NP complement, and so would be conventionally classified as “adverbs” in both Hausa and English, because prepositions require complements. Modern accounts, however, extend the prepositional category to consist of a head preposition with or without an overt complement, i.e., transitive and intransitive prepositions respectively. Once again there is a functional/syntactic analogy with verbs – they behave no differently from (dual-transitivity) verbs occurring with or without a complement, e.g.,

- 35.a. *sun kaamàa shi*
 3PL.PFV grab 3MSG
 ‘they grabbed him’

Cf. intransitive:

- b. *àbîn dà ya kaamàa dàgà nân...*
 thing.the REL 3MSG.PFV reach from here
 ‘what reaches from here...’

- 36.a. *kanàa cîn goořò?*
 2MSG.IPFV eating.of kolanut
 ‘do you eat kolanut?’

Cf. intransitive:

- b. *kàasuwaa tanàa cîi*
 market 3FSG.IPFV eating
 ‘the market is in full swing (lit. ‘is eating’)

Occurrence of the same verb **kaamàa** in both the transitive (35a) and intransitive (35b) clauses is not a valid reason for proposing that the two words belong to different classes—both belong to the same primary category *verb*. The same argument applies to verbonominal transitive **cîn** (36a) and intransitive **cîi** (36b). The difference relates to the complements they take. By the same token, there is no motivation to treat (33) **yanàa [cikin ajii]** ‘he’s in(side) the classroom’ and (34) **yanàa [ciki]** ‘he’s inside’ any differently, they are transitive and intransitive prepositions respectively.

Another problem for any traditional analysis is that prototypical adverbs are defined as *semantically modifying verbs*. But the static locative **ciki** ‘inside’ in (34) **yanàa [ciki Ø]** ‘he’s inside’, for example, hardly meets the traditional definition of ‘adverb’ – it is the locative complement of the inflectional [3ms.impfv] auxiliary element, and, more seriously, there is *in fact no verb present* to be modified. Finally, in (37) variation in the type of complement would not require us to attach distinct grammatical labels to ‘buy’ in (a) and (b). This would be a pointless complication.

37. a. *kin* *sàyi* *shìnkaafaa?* (= transitive verb with NP COMP)
 2FSG.PFV buy rice
 ‘did you buy some rice?’

Cf.

b. *kin* *sàyaa Ø?* (= same lexical verb with ellipted COMP)
 2FSG.PFV buy
 ‘did you buy (some rice)?’

6. Summary: different prepositions take different complements

Taken together, the above evidence — mainly occurrence in parallel morpho-syntactic environments — represents a perhaps radical but compelling case for migrating many traditional conjunctions and adverbs into an expanded prepositional category, thereby tidying up one area of the grammar. As demonstrated above, many Hausa prepositions, as heads of phrases, govern different types of dependents, ranging from clauses and NPs to zero complementation. It is hoped that the linguistic data and analyses might encourage other linguists to think in fresh ways about the category preposition and its membership.

References

- Bagari, Dauda M. 1987. *Hausa Subordinate Adverbial Clauses: Syntax and Semantics*. Rabat [Morocco]: El Maarif Al Jadida.
- Caron, Bernard. 1987. Description d’un parler haoussa de l’Ader (République du Niger). Thèse de Doctorat d’État, Université de Paris 7.
- Emonds, Joseph E. 1972. “Evidence that indirect object movement is a structure-preserving rule”. *Foundations of Language* 8, 546-61.
- Haruna, Andrew. 2003. *A Grammatical Outline of GùrdùN/Gùrùntùm (Southern Bauchi, Nigeria)*. Cologne: Rüdiger Köppe.
- Jackendoff, Ray S. 1973. “The base rules for prepositional phrases”. In: Stephen R. Anderson & Paul Kiparsky (eds.). *A Festschrift for Morris Halle*. New York: Holt, Rinehart & Winston, 345-356.
- Jaggar, Philip J. 1992. *An Advanced Hausa Reader with Grammatical Notes and Exercises*. London: SOAS.
- Jaggar, Philip J. 2001. *Hausa*. Amsterdam & Philadelphia: Benjamins.
- Jaggar, Philip J. 2010. “Relational DÀ as a preposition, coordinator, and subordinator in Hausa: Polysemy or homophony?” In: Georg Ziegelmeyer & Norbert Cyffer (eds.). *Aspects of Co- and Subordination: Case studies from African, Slavonic and Turkic Languages*. Cologne: Rüdiger Köppe, 61-77.
- Jespersen, Otto. 1909-49. *Modern English Grammar on Historical Principles* (7 vols). London: Allen & Unwin.
- Jespersen, Otto. 1924. *The Philosophy of Grammar*. London: Allen & Unwin.

- Jungrauthmayr, Herrmann, in collaboration with Njeno Andirya Galadima and Ulrich Kleinewillinghöfer. 1991. *A Dictionary of the Tangale Language (Kaltungo, Northern Nigeria) with a Grammatical Introduction* [Sprache und Oralität in Afrika, Vol. 12]. Berlin: Dietrich Reimer.
- Kraft, Charles H. 1963. *A Study of Hausa Syntax*. Vol. 2. Hartford: Hartford Seminary Foundation.
- Leger, Rudolf. 1994. *Eine Grammatik der Kwami-Sprache (Nordnigeria)*. Cologne: Rüdiger Köppe.
- Newman, Paul. 2000. *The Hausa Language: An Encyclopedic Reference Grammar*. New Haven, Yale University Press.
- Newman, Paul. 2007. *Hausa-English Dictionary*. New Haven: Yale University Press.
- Newman, Paul & Roxana M. Newman. 1977. *Modern Hausa-English Dictionary (Sabon K'amus na Hausa zuwa Turanci)*. Ibadan & Zaria: Oxford University Press (Nigeria).
- Pawlak, Nina. 2007. "The boundaries of the matrix clause: multiverb constructions in Hausa". *Lingua Posnaniensis* 49: 97–108.
- Pullum, Geoffrey K. 2009. "Lexical categorization in English dictionaries and traditional grammars". *Zeitschrift für Anglistik und Amerikanistik* 57(3), 255–273.
- Pullum, Geoffrey K. & Rodney Huddleston. 2002. "Prepositions and preposition phrases". Chapter 7 of Rodney Huddleston & Geoffrey K. Pullum *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Schuh, Russell. G. 1981. *A Dictionary of Ngizim*. Berkeley & Los Angeles: University of California Press.
- Schuh, Russell. G. 1998. *A Grammar of Miya*. Berkeley & Los Angeles: University of California Press.
- Wolff, H. Ekkehard. 1993. *Referenzgrammatik des Hausa*. Münster & Hamburg: Lit.

Dorothee Beermann

ORCID 0000-0002-8906-5235

NTNU, Trondheim, Department of Language and Literature

Polytext AS

Lars Hellan

ORCID 0000-0003-3043-120X

NTNU, Trondheim, Department of Language and Literature

Polytext AS

Enhancing grammar and valence resources for Akan and Ga

Abstract

We present a case study in valence comparison between closely related Kwa languages, assessing frames and meanings of the verb *ba* ('come') in Akan with a homophonous corresponding item in Ga. The discussion draws on the Akan dictionary (Christaller 1881), a Ga valence dictionary based on (Dakubu 2009), and an online annotated corpus of Akan hosted in TypeCraft (Beermann & Mihaylov 2014). With a view to the possibility of making use of resources for one language in the development of resources for another, we demonstrate how digital resources and linguistic specifications can inform each other.

Keywords: Akan, Ga, digital resources, lexicon, valence, corpus annotation

1. Introduction¹

Akan and Ga² are closely related Kwa languages spoken in Ghana. Ga with about 745,000 speakers in 2013³ is a small language spoken in the area of the capital Accra, while Akan with over 9 million L1 speakers is a widely spoken language in Ghana and along Africa's west coast. Akan is one of the official languages of Ghana alongside English and is understood widely throughout the country; most speakers of Ga thus also speak Akan but not vice versa. Akan is used extensively as written language in literature and in education, and in the media including the Internet, while for Ga this is much less so.

The languages are known for their many types of serial verb constructions (SVCs) and other multi-verb constructions, and the notion of valence is potentially harder to systematize for such languages than for languages with predominantly mono-verbal sentence patterns.⁴

¹ We are grateful to the editors of the volume for their careful commenting, and to Abigail Agyeiwaa for helping us with the Akan data.

² Akan: ISO-639-3 [aka], Ga: ISO-639-3 [gaa]

³ <https://www.ethnologue.com/country/GH/languages>

⁴ Cf. Beermann & Hellan (2018).

This paper is a preliminary to an investigation into valence of Kwa languages.⁵ For the present study we take advantage of Christaller’s seminal “Dictionary of the Asante and Fante Languages called Tshi” (Christaller 1881), and of digitally encoded valence lexicons for Ga (Dakubu 2010; 2011). Among available digital resources are: a labeled corpus of 98 000 word tokens for Akan corresponding to in depth annotated 7 941 phrases and a small corpus (461 sentences) of Ga. Given the close relationship between the two languages, they could be a ground for a multilingual approach to resource development for these two languages: the grammars are largely similar and so are many of the words and phrases; but when it comes to a particular word or a particular construction, there is no a priori way of deciding what is identical across the languages. Our discussion will focus on one single verb in each language and its syntactic and semantic frames; this verb has the form *ba* in both languages and has been assigned the English gloss ‘come’ by the respective scholars of each language. We consider it as a sensible probe for investigating the possibilities of cross-linguistic comparison and resource development.

The paper is structured as follows: in sections 2 and 3 we describe the existing digital resources and analytic and representational tools. In section 4 we render information on *ba* in Akan from Christaller (1881). Section 5.1 describes information in the Ga valence lexicon concerning *ba*, section 5.2 describes information in the Akan corpus concerning *ba*, and 5.3 renders results from interviews with Akan speakers. Section 6 summarizes the main findings and assesses aspects of contrastive valence studies in a setting like this.

2. Resources for Akan⁶

Our Akan corpus has been developed in TypeCraft (TC),⁷ a linguistic service for the management and annotation of Interlinear Glossed Text (IGT) from lesser resourced languages. The tool features:

- Support for the creation and management of larger corpora and user-space texts.
- An editor for indepth morpheme and word level annotation.
- Support for the annotation of Multi-word expressions and discourse senses.
- Access to over 2000 individual text-instances across 145 mainly less resourced languages.
- Export-Import facilities for ELAN, eaf, csv and XML data.
- Export of examples in the Leipzig Glossing format to all main text editors.

Launched in 2005, TypeCraft was the first online interlinear glossing tool, which allowed for real-time sharing of small corpora, enabling users to annotate, edit, and collaborate on the creation of linguistic data wherever they were, with a tool suitable for any

⁵ For studies on valence in other West African languages, cf. Creissels (2015) and Schaefer & Egbokhare (2015).

⁶ In the present outline of Akan resources, we restrict ourselves to those digital resources that are available for our tools and methodologies. To our knowledge, lexicology works like Christaller (1881) and Anyidoho (2006) are not available for these methodologies, nor are there to our knowledge accessible corpora of Akan other than those to be discussed.

⁷ Cf. Beermann & Mihaylov (2014) and Beermann (2015).

user. Glossed data thus was no longer locked to a single computer or a single article but could be accessed at any time through the TypeCraft's online portal. In this respect TC still carries a certain advantage over linguistic desktop systems, such as ELAN.

TypeCraft's Akan corpus has been actively developed in recent years, and with over 80 000 sentences is the largest of our African corpora⁸, while Ga with 461 sentences is considerably smaller.

In linguistic publications, examples annotated following the Leipzig Glossing Rules (LGR) are seen as a standard. Examples are glossed on three tiers showing the morphological break-up of words, followed by a line with morphosyntactic and English glosses, and a third tier with a free translation mostly into English. Although LGR features some Parts of Speech (POS) labels, systematic POS-tagging is not pursued as opposed to what is standard in Natural Language Processing (NLP). The lack of POS tagging is in many respects a disadvantage not only for the automatic processing of linguistic resources, but also for their linguistic analysis, and TypeCraft therefore has chosen to assign a separate tier to POS tagging. The system supports several closed annotation sets, and along with the standard morphosyntactic glosses, valence tags and basic Discourse Sense tags are available.

The Akan corpus referred to here has been created by native Akan speakers for seminar work and graduate projects in linguistics. Individual projects on TypeCraft do not reflect a shared annotation strategy, as individual users of TypeCraft pursue different linguistic interests, and are free to use the system as suits their goals best. The Akan corpus curated by the TypeCraft group is in this respect different from the user-driven TypeCraft corpus.

A curation effort was conducted between 2015 and 2018 (Beermann et al. 2018; Van Dommelen & Beermann 2019) for part of the TypeCraft Akan corpus. As a result, this TypeCraft Akan corpus has a higher analytic depth (more features) and an increased consistency, which distinguishes it from some of the privately owned Akan data on TypeCraft.

3. Resources for Ga

The digital lexicon (Dakubu 2010) is an amended version of a Ga Toolbox lexicon project holding data for the general-purpose dictionary (Dakubu 2009). *Ga Verbs and their constructions* (Dakubu 2011) is a free-standing linguistic monograph. Here we will mainly refer to the former work. This resource consists of 80,000 lines of code with 7080 entries, 5014 of which are for nouns and 935 for verbs. Table 1 illustrates an entry in the valence enriched format for the verb *ba* meaning 'come', where the valence codes are written into the lexical entry following the general field style of Toolbox, here as the fields \s1, \s2, \s4, \s6:

```
\x ba
\hm 1
\ph ba
\ps verb annotated
\pdl neg. imperf
\pdv baaa
```

⁸ Here we also count all our Akan data and also texts that have not been annotated.

```

\pdl neg. perf
\pdv bako
\pdl neg. fut
\pdv baŋ
\pdl imper
\pdv bá
\ge come
\sl1 v-
\sl2 intr-
\sl4 suAg-
\sl6 MOTIONDIRECTED
\xv E-ba oya
\xg 3S.AOR-come quickly
\xe he came quickly.
\xv È-bá-aa
\xg 3S-come-NEG.IPFV
\xe he didn't come; he didn't measure up (to a task).
\dt 27/Dec/2009

```

Table 1. Example of Ga Toolbox entry enriched with valence information

The fields marked \pdl-\pdv represent inflectional information of the lexeme as follows: negative imperfective: *baaa*; negative perfective: *bako*; negative future: *baŋ*; imperative: *bá*. \ge gives the English gloss *come*. The fields \xe, \xg, \xv together constitute a standard linguistic glossing with \xv as a word-and-morph break-up, \xg as morphological and English gloss, and \xe as a free English translation.⁹

A verb with more than one valence frame has one entry specified per frame; thus the verb *ba* ‘come’, for instance, is represented by 18 different entries in this edition of the Toolbox file. In this way, 547 verb lexemes from the original file are represented through 2006 entries altogether.

The valence specification follows principles and formalization laid out in Hellan & Dakubu (2010),¹⁰ the *Construction Labeling (CL)* system. In this formalization the information located in the fields \sl1, \sl2, \sl4, \sl6 in Table 1 can be rendered in the more compact form given in (1), to be read as ‘a verb-headed intransitive syntactic frame where the subject carries an agent role and the situation expressed belongs to the type ‘MOTIONDIRECTED’.¹¹

(1) v-intr-suAg-MOTIONDIRECTED

⁹ With such IGTs illustrating verbs and smaller phrases illustrating nouns and other POS, these specifications in the Toolbox file constitute a large corpus also for accommodation in TC, given a conversion script which yet remains to be implemented, potentially applicable to Toolbox files for other languages as well.

¹⁰ Also see Dakubu & Hellan (2017).

¹¹ A feature structure reorganization used by Dakubu (2010) is done by Bruland (2011), accessible at (this feature structure lexicon is part of a computational Ga grammar Hellan & Beermann (2014).

The semantic specification in (1) consists of two parts, *semantic role* as exemplified by ‘suAg’ and *situation type* as exemplified by ‘MOTIONDIRECTED’, the latter out of a total inventory of about 130 situation types.¹² Situation types are word-independent, and as a tendency, less specific than word meanings. Thus, word level instantiations of a given situation type may involve one word in one language, sequences of many words in another (cf. discussion in Hellan (2019)). The most frequently used situation types in the Ga valence lexicon entries are given in Appendix 2. The number of individual verb meanings in this lexicon is more than 500, in contrast, attesting to the difference between situation type and verb meaning.

The classification using all the parameters recognizes about 100 construction types, which for mono-verbal constructions could also be seen as valence types. A small TC-corpus illustrates these construction types.¹³

With a set of 2000 entries classified by strings like (1), the valence notation allows one to investigate the frequency of frames used relatively to these frames, correspondences between syntactic and semantic structure, the clustering of certain valence types for sets of verbs, and more.¹⁴ For instance, it appears that the lexically by far most common valence frame is *transitive*, relatively few verbs occur only in *intransitive* constructions, and most verbs that occur ditransitively can also occur with only one object. In this vein,¹⁵ Appendix 1 shows the number of verbs which can occur in various valence frames, where also occurrence in a serial verb construction (SVC) is included, some in only one type of frame, others in more than one.

Ga makes little use of prepositions and adjectives, so that constructions involving nouns and verbs may be seen as playing a relatively large role, the latter, for instance, through multiverb expressions subsuming not only SVCs but also *Extended Verb Complexes (EVCs)*, which are combinations of a main verb and verbs with a predominantly functional nature, like inducing deixis, path aspect or other features. When such functional verbs precede the main verb, they are often referred to as *preverbs*, but since they can also occur post-verbally, we prefer the order-neutral notion EVC for the whole complex and “functional satellite verb” for the functional verbs; following common usage, we also refer to them as “preverb” and “postverb”, respectively. Still a further type of multiverb expressions are *Verbid Constructions (ViD)*, where verb phrases play the role of adverbials; (for description and discussion of these constructions, see e.g. Dakubu (2004a; b; 2008) and Dakubu et al. (2007)).

¹² See Dakubu (2011) and Hellan & Dakubu (2010). The frame types used in FrameNet (<https://framenet.icsi.berkeley.edu/fndrupal/>), where ‘frame’ could be conceived as corresponding to what we here call situation type, were found to be too English-biased to immediately serve in this lexicon. See Appendix 2 for a sample of situation type labels.

¹³ See https://typecraft.org/tc2wiki/Ga_Valence_Profile. The data are searchable, so that a search for, e.g., the constructional factor *obPostp* (‘object is a postposition’) yields an array of urls for the sentences instantiating the factor.

¹⁴ Studies from this perspective, often attributed to Levin (1993) as a starting point, and Malchukov & Comrie (2015) are also applied to Ga (Dakubu & Hellan 2016).

¹⁵ Files with entries under each combination are accessible at https://typecraft.org/tc2wiki/Ga_Valence_Profile in a zipped version, with meaning and examples obtained at the link [Ga_verb_dictionary_for_digital_processing](https://typecraft.org/tc2wiki/Ga_Valence_Profile).

The specifications of the Toolbox valence lexicon are also used for classification in a Ga lexicon with the 2000 verb entries at the online (4-language) valence lexicon *MultiVal*.¹⁶ The lexicon is also used in a computational grammar of Ga based on the HPSG¹⁷ framework.¹⁸ These are examples of how information, once digitally encoded, can be recast in other formats and used for other purposes.

As already said, the present discussion will focus on the verb *ba* ‘come’ in both languages. We return in section 5 to a view of what the Ga lexicon provides for the item. We first look at *ba* in Akan.

4. Christaller about *ba*

Christaller’s entry for the verb *ba*, to which he assigns the basic meaning ‘come’, densely lists syntagma and senses associated with the lexeme (Figure 1).

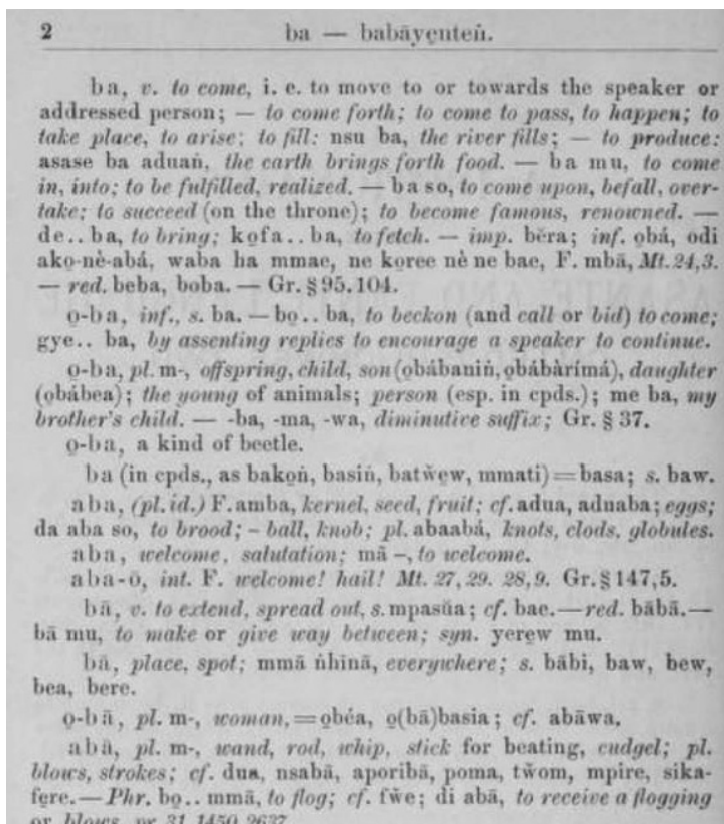


Figure 1 The entry for *ba* in Christaller’s dictionary of Twi

¹⁶ Cf. Hellan et al. (2014).

¹⁷ Head-driven Phrase Structure Grammar (Pollard & Sag 1994).

¹⁸ Cf. Dakubu et al. (2007) and Hellan & Beermann (2014).

According to this entry, the intransitive syntactic frame of *ba* describes situations of *directed movement* and features a *mover* subject. Intransitive *ba* may – according to Christaller – also describe non-directed motion in the sense of a *happening*. Christaller further notes that *ba* is also used as a ventive marker where it expresses movement to a deictic center (e.g., the speaker, but not exclusively). We will see that this use is realised in an EVC with *ba* preceding the main verb.

An intransitive frame with *ba* may also express a directed movement along the vertical axis giving rise to the reading ‘to fill’ or ‘to rise’. As an example Christaller mentions *nsu ba* which he translates as ‘the river fills’.

According to Christaller *ba* may enter a transitive frame in two ways. The first construction expresses a situation of becoming; this use of *ba* is also attested for Ga. Christaller gives the following example: *Asase ba aduan* ‘the earth brings forth food’.¹⁹ The second transitive use is connected with the expression of Endpoint Motion. In Christaller’s examples the verb is shown in construction with a relational noun where it is left open what may intervene. Left open is also the question whether the relational noun builds an embedded phrase with the word that directly precedes it. Thus, Christaller translates *ba ... mu* as ‘to come in or into’, or as ‘to be fulfilled or to be realised’. *Ba ... so* is translated as ‘to come upon, befall, overtake or succeed’. The relational nouns *mu* and *so* denote an inside place or a place on top, respectively. Both constructions are attested also in Ga and will be discussed in more detail in section 5.

Next to taking an intransitive and a transitive frame and its role in EVCs, *ba* also plays a special role in SVCs of the type that Christaller calls “essential” combinations and which we classify as *integrated SVCs* (ISVC) following Osam (1994). These are combinations where two verbs form a meaning which does not seem composed from their respective meanings.²⁰ In the contrasting type of SVC, by Christaller called “accidental” and here referred to as *chaining SVCs*, the verbs retain their status as independent predicates. Christaller gives the following examples of SVCs of the respective categories – full examples will be discussed later:

Integrated SVC	
de ... ba	to bring
ko ...fa...ba	to fetch
Chaining SVC	
bo ...ba	to call someone to come
gye ... ba	to encourage a speaker to continue

¹⁹ On p. 2 in the 2013 edition of Christaller (1881).

²⁰ Since *de ... ba* means ‘bring’, it could be argued that the expression is transparent, as *de* means ‘take’, ‘taking’ and ‘coming’ being components of ‘bringing’. Such transparency does not hold for all ISVCs, however.

5. A comparative analysis of valence, constructions and situation types involving *ba*

5.1. *Ba* in the Ga valence lexicon

From the Ga valence lexicon we will consider the 18 entries assigned to *ba*, illustrated below by their construction types and an annotated example reconstructed from the Toolbox file (*ba_1* corresponding to Table 1). In each case the gloss entered in the original version is ‘come’ and is thus not repeated; comments on notation and analysis are offered throughout.

- (2a) *ba_1* : v-intr-suAg-MOTIONDIRECTED
 È-bá-aa
 3SG-come-NEG.IPFV
 ‘he didn’t come’
- (2b) *ba_2* : evSuAg-vintr-pv1obTh-MOTIONDIRECTED
 E-kɛ kakla lɛ ba
 3SG -move knife DEF AOR.come
 ‘she brought the knife.’

Label specification: ‘evSuAg’ here stands for “extended verb complex with shared subject as Agent”, ‘vintr’ means that the main verb is intransitive, and ‘pv1obTh’ means that preverb 1 (the label here used for the functional satellite verb when preceding the main verb) has a theme object, which is to say that the preverb *kɛ* is transitive.²¹ A bit unclear is why the situation type of this construction is given as MOTIONDIRECTED and not rather, for instance, TRANSPORT.

- (2c) *ba_3* : v-tr-obPostp-suAg_obLoc-MOTIONDIRECTED
 E-ba tsu mli
 3SG-AOR-come room inside
 ‘he entered the room .’

The label ‘obPostp’ means “object is a postpositional phrase”, a label reflecting the tradition of calling relational nouns in this kind of role “postpositions”, whereas in calling *ba* here transitive, the author in effect counts *tsu mli* as a noun phrase.

- (2d) *ba_4* : v-tr-obPostp-suTh_obLoc-PHENOM
 Lamo lɛ ba mli
 dream DEF AOR.come inside
 ‘the dream came true.’

²¹ A configuration described by Dakubu (2004a).

- (2e) *ba_5* : v-trComp-suExpl-obPostp_compDECLcmp-obLoc-PHENOM
 E-ba mli akε.
 3SG.AOR-come inside COMP (S)
 ‘it happened that..’

The example is here only schematically indicated, given with the author’s comment: “precedes the main verb, indicating event movement towards the speaker”. *Mli* is here counted as syntactic object, and the event as such is expressed by a declarative complement with *akε* as complementizer.

- (2f) *ba_6* : v-tr-suAg_obEndpt-MOTIONDIRECTED
 E-ba shi
 3SG.AOR-come down
 ‘he came down; he bowed’

The item *shi* is counted as a noun.

- (2g) *ba_7* : v-tr-suTh_obEndpt-REDUCTION
 Nyɔŋmɔ lε é-ba shi
 sky DEF PERF-come down
 ‘the rain has let up’.

- (2h) *ba_8* : v-ditr-suAg_obTh_ob2Endpt-MOTIONDIRECTED
 O-ba lε shi
 2SG.AOR-come 3S down
 ‘You lowered it.’

The 3sg pronoun *lε* is here counted as first object and *shi* as second object, thus a double object construction which, however, has little association with what one standardly would call “ditransitive”; *lε* seems counted as direct object of *ba* whereas *shi* is more like an independent endpoint expression.

- (2i) *ba_9* : v-tr-suPostp-suLocus_obLoc-REDUCTION:
 E-naa ba shi
 3SG.POSS-edge AOR.come down
 ‘it lessened, was reduced.’

E-naa ‘its edge’ is a subject phrase headed by a relational noun *naa*; *shi* is counted as object, like above.

- (2j) *ba_10* : v-tr-obPostp-suAg_obLocus-EXHIBPROPTY
 É-ba nɔ
 3SG.PERF-come PRO.POSS on
 ‘she has become very disrespectful.’

- (2k) ba_11 : v-intr-suTh-PHENOM
 Aduawa le ba
 seed DEF AOR.come
 ‘the seed germinated.’
- (2l) ba_12 : v-tr-obBPobSpec-suTh_obLoc-SENSATION
 Kelewele e-ba mi-tsine
 kelewele PERF-come 1SG.POSS-longing
 ‘I’m craving some *kelewele*.’

The object is analyzed as body-part of the specifier in the object phrase, thus, a “longing” as part of me.

- (2m) ba_13 : v-tr-obPostp-suTh_obLoc-INCREMRESULT-PROPTY²²
 Daa le e-ba e-hie
 drink DEF PERF-come 3SG.POSS-face
 ‘he is tipsy.’
- (2n) ba_14 : v-ditr-obPostp-suAg_obLocus_ob2Loc-REFLEXIVE-EXHIBPROPTY
 E-ba e-he shi
 3SG.AOR-come 3SG.POSS-self down
 ‘he humbled himself.’
- (2o) ba_15 : v-tr-suIDobSpec-suAg_obLoc:
 E-ba-a e-jeŋ jogbaŋŋ
 3SG-come-HAB 3S.POSS-world well
 ‘she behaves well.’

The label ‘suIDobSpec’ means “subject is identical to the specifier of the object *e-jeŋ*”, so ‘she comes her world’; *jogbaŋŋ* ‘well’ is an adverb.

- (2p) ba_16 : v-tr-suEmitter_obTh-EMISSION
 Shikpɔŋ nɛɛ ba-a abele jogbaŋŋ
 land this come-HAB corn well
 ‘this land produces corn well.’
- (2r) ba_17 : evSuExp-pvI intr-vtr-obPrept-SENSATION
 E-ba-na mi
 3SG-INGR.AOR-see 1S
 ‘she came and saw me.’

²² For ‘property resulting from an incremental process’ (see term explanations at the end).

In this and the next example *ba* functions as preverb.

- (2s) ba_18 : evSuAg-vtr-pv1obDrop-obTh-SENSATION
 E-baa-kɔ
 3SG-INGR.FUT-take PRO
 ‘he will come take it.’

Both the main verb *kɔ* ‘take’ and the preverb *ba* are here annotated as transitive, with a dropped object annotated for the preverb, whereas the main clause object is not explicitly marked as dropped, as notation would dictate, although it appears to be. When *ba* occurs as a preverb its function is otherwise mainly to indicate speaker-orientation, hence it is not easy to make sense of a dropped object for *ba* here, whereas for *kɔ* it does make sense.

To comment on these examples further, five of the above entries are labelled as expressing the situation type MOTIONDIRECTED.²³ Three entries are classified as having a syntactically intransitive frame, two of them with *ba* as preverb, whereas ten instances of *ba* are classified as having syntactically transitive frames, and two entries are labelled as having a syntactically ditransitive frame. In the transitive category is for instance ba_3 with the example *Eba tsu mli* (3S-AOR-come room inside ‘he entered the room’), where *tsu mli* denotes the endpoint of the directed motion and is analyzed as a syntactic object. The examples in ba_6 and ba_8 show a similar use of nominals expressing endpoints, now with the noun *shi* (glossed as ‘down’, but with “noun” as part of speech). In ba_6 *shi* is treated as the single object whereas in ba_8 the items *le* and *shi* are identified as the first and the second object respectively, and accordingly *ba* is classified as a double object construction.

For a verb whose standard gloss is ‘come’, the proliferation of transitive occurrences is perhaps somewhat surprising. For one thing, one may wonder if the basic meaning of the item *ba* is more like ‘approach’ or ‘reach’, which to an English *Sprachgefühl* would make the taking of objects more plausible. Another issue is the status of relational objects and the possibility that they are perhaps rather functionally on a par with prepositions and adverbs, whereby supposed objects with *mli* and *shi* would be after all more like adverbs or prepositions. The choices of analysis shown in (2a-s) seem to be consistently carried through the larger lexicon, and we are not in a position to discuss either issue, but in the possible context of initiating a similar endeavour for Akan, and perhaps for the two languages in tandem, both issues as relating to the notion of transitivity deserve, of course, a general discussion prior to a larger-scale effort.

A general issue is that when *ba* occurs as last verb in a VV sequence of the type ISVC or EVC, its apparent valence could be influenced by the collective status of the two Vs. In (2) this constellation is instantiated only in ba_2 where *ba* follows the transitive preverb *ke*; here, however, *ba* itself is assigned the syntactic frame *intransitive* despite the transitive ‘neighbourhood’. Still, we will encounter the issue when considering the Akan data, cf. 5.2.

²³ In the whole lexicon 55 entries have the situation type MOTIONDIRECTED, the other entries with this type are glossed for lexical meaning as ‘travel’, ‘start’, ‘turn around’, ‘lucky’, ‘proceed’, ‘climb’, ‘land’, ‘run’, ‘visit’, ‘arrive at’, ‘forget/leave’, ‘depart’, ‘go’, ‘bother about’, ‘push (away)’, ‘go before’, ‘strike’, ‘trail’. Cf. https://typcraft.org/tc2/wiki/Ga_Valence_Profile.

The Ga lexicon is unique as a digital Kwa valence dictionary. It offers a fine-grained and consistent labeling of verb dependents and of multi-verb constructions which are characteristic for this language family. It thus – in conjunction with Dakubu (2011) – can serve as a basic reference point in explorations of valence in Kwa. If used in tandem with an investigation of Akan, this resource of Ga can, of course, in turn be further developed.

5.2. *Ba* in the Akan corpus

In our corpus of Akan we find 155 sentences which contain *ba*. The corpus is too small for quantitative studies, nevertheless, it allows us to identify intransitive, transitive, EVC and serialisation uses of *ba*. Since some of the data is from conversations, transcribed descriptions of video clips or spoken code-switching Akan-English,²⁴ it also allows us a more realistic look at the internal structure of transitive *ba* constructions and serialisations.

In the corpus, *ba* occurs often in construction with a pronominal place expression, exemplified in (3) (in a format exported from TypeCraft). Whether this construction is transitive with the pronoun in an object-like function, or rather with the pronoun as an adverbial, is most often not indicated in the annotation (although in (3) it is indicated as adverb), and a decision will have to address the issues concerning transitivity mentioned above for *ba* in Ga.

- (3) Papa akɛsɛɛ nsia yi baa ha
 papa a kɛsɛɛ nsia yi ba a ha
man.SBJ PL big six this.DEF come PAST here
 N ADJ NUM DEM V ADVplc
 ‘These big men came here.’

In the example (4), elicited in interview,

- (4) ɔba adwummam hɔ
 3SG-come workplace there
 ‘He/she comes to the workplace’

we do not have a locative noun that is a noun in construction with a relational noun, but a simple noun phrase, and here transitivity is only licensed when the noun is followed by a place adverb like *hɔ*.

We have found one case where *ba* is in direct construction with a proper noun as its direct object, but in this case the noun is semantically a locative noun, and thus these examples together might indicate that transitivity of *ba* is only licensed when the object is a locative expression.

²⁴ Based amongst others on Brobbey (2015).

- (5) *Ernest kɔ Ghana kɔ waree de ne yere baa Norway*
 NP+NP-directedMotion-----
 ernest kɔ ɔ ghana kɔ ware de ne yere ba a
 SBJ go PAST go take 3SG wife.SBJ come PAST
 Np V N V1 V2 V3 PNposs N V4
 norway
 Np
 ‘Ernest went to Ghana to marry and brought his wife to Norway’

A further example of ENDPOINTMOTION, with an NP (possibly analysed as PostP) as endpoint expression, is given in (6), also parallel with Ga examples:

- (6) A-n-kye biara na ɔ-pue ba-a nsuo no ani
 PAST-NEG-last at all NARR 3SG-emerge come-PAST water DEF top
 V ADV PRT V V N DET N
 ‘It didn’t take a long time and he came to the surface of the water’

In (6) the function of *ba* may be as a postverb to induce a PATH construal, which would make (6) an EVC; alternatively this is an ISVC with *pue* and *ba* as merged predicates sharing the locative noun as their object.

The use of *ba* as preverb is exemplified below, within the last VP of a Chaining SVC. The chaining is realised by the juxtaposed verbs *ba* and *sere* sharing a subject. *Sere* in the second conjunct of the serialisation is preceded again by *ba* here realised as *be*²⁵. *This is the ventive use of ba* in an SVC, a use already noted by Christaller, and not unlike the use of *ba* as preverb in Ga.

- (7) ɔyɛ n’adwene sɛ ɔbeba abesere no
 ɔ yɛ ɛ n’ adwene sɛ ɔ be ba
 3SG make PAST POSS.3SG mind that 3SG FUT come
 V N CONJ V1
 a be sere no
 CONS VEN beg 3SG
 V2 PN
 ‘He made up his mind that he will come and beg him’

An example unlike any of those discussed for *ba* in Ga but mentioned in the comments on Christaller, is one of *ba* in an Integrated SVC as given in (8), where *ba* and the preceding verb in the VV do *not* have a meaning that they would carry also in a mono-clausal environment.²⁶

²⁵ A view of allomorphic variation of *ba* in such combinations is given in Appendix 3.

²⁶ In the TypeCraft Akan corpus, a search (April 01.2019) for the translation string ‘bring’ yields 28 phrases, of which half use *de...ba* as expression for ‘bring’, and the other half uses *fa* ‘take’ together with a verb for ‘give’ such as *ma*, the latter arguably a further instance of a constructional expression of ‘bring’.

- (8) am no so a na wo de bac.
 meat DEF how.much REL FOC 2SG use come-PAST
 N DET ADJ PRT PRT PN V V
 'how much game did you bring.'
 (<https://typecraft.org/tc2/ntceditor.html#3622,453767>)

Here *de* and *ba* together express 'bring', as indicated in the free translation, a concept which reflects a combination of the meanings 'take/use' and 'come' indicated in the gloss line. The sequence *de ... ma* is more like a constructional unit, whose meaning must be taken out of a space of notions that are ascribed to sentences, thus situation types.

5.3. Interviews

Aside from the consultation of Christaller and the Akan corpus, combined with common knowledge of the Akan linguistic literature, we have also conducted interviews with Akan speakers (students of English linguistics), applying the list of constructions in (2a-s) as a tentative probe into what possible Akan constructions are. Limitations in time and funding do not make it possible to conclude how far such a procedure can be used, but our results indicate that nearly all the patterns in (2a-s) can be replicated in Akan with the use of *ba* in a similar role as *ba* has in the Ga examples. Here we only mention two examples:

- (9a) 9a) ba
 3SG-go_down (DEIXIS|PATH)
 'He/she goes down'
- (9b) ɔ-foro ba
 3SG-go_up (DEIXIS|PATH)
 'He/she gets/moves up'

Also in (9a-b) the function of *ba* seems to be that of a postverb in an EVC construction, indicating either deixis or path – which of these will have to be decided in context.

6. Conclusions

In addition to summarizing the main findings already stated, we would like to bring up a few points relating to linguistic resources, and a few points of a more theoretical nature.

As for main findings, with a status for *ba* in Akan as found in Christaller (1881), we have excerpted entries for *ba* in the Ga valence dictionary (Dakubu 2010), with a number of 18 different environments for *ba* viewed in terms of syntactic environment and situation type, all with 'come' as a supposed "lexical" meaning, and then have seen in our Akan corpus and in interviews whether a similar array of uses can be stated for the item *ba* in Akan. There is correspondence, but the inspection of the Akan corpus seems to suggest that the role of *ba* in VVs expressing TRANSPORT with constructional glosses such as 'bring', 'fetch', 'get', etc. may be more prominent than it is suggested by the Ga lexicon.

Otherwise *ba* both in Ga and Akan can be used in extended verb complexes, where *ba* contributes to either deixis or path information relative to the main predicate. In this function *ba* can either be a pre- or a postverb.

In both languages its role in ENDPOINTMOTION constructions faces the same analytic choices regarding transitivity, i.e., whether ‘come’ is at all an adequate rendering of the word in isolation, and to what extent the combination of *ba* with a place denoting noun is a syntactically transitive construction rather than an intransitive construction with an adverbial or postpositional phrase. The latter issue concerns the full range of motion verbs in both languages, and thus addresses principled aspects of the grammars of both languages.

Investigations like these should ground themselves methodologically as broadly as possible. Digital resources are represented here by a digital annotated corpus and by a digitally represented valence lexicon; not digital in format is Christaller’s lexicon. Neither type of digital resource is perhaps a *must* for investigating valence, but either type is helpful when one has it. For both types of digital resources it is essential that annotations and specifications are kept consistent and phrased in terms that assure transparency of the analytic choices that have informed the annotation.

In his preface, Tesnière (1959) expresses regret that time did not allow him to investigate the notions of grammatical relations and valence in languages outside of the Indo-European space that served as basis for his system. Many features of the Kwa languages investigated here, not only with regard to the many types of multiverb expressions but also to the features of the nominal system (some illustrated above), constitute interesting challenges and grounds for development relative to this system (and therewith to the bulk of contemporary syntactic systems like LFG²⁷, HPSG, Universal Dependency Grammar (UDC)²⁸, and many more). Within valence studies in turn, *contrastive* valence studies have become a topic of interest (Malchukov & Comrie 2015; Hellan et al. 2017), and our comparison of valence in Akan and Ga touches on this topic. A main challenge for such studies is the circumstance that between two languages one cannot predict in what respect they may have features identical or similar, no matter how close the languages are. Given the resource asymmetry between Ga and Akan (one with a large valence dictionary but very little corpora, the other with no valence dictionary but with developed morphological corpora), one might, for instance, envisage scenarios where the more developed aspects of one resource can be more or less automatically transferred to the other, for example, the build-up of an Akan valence lexicon being modeled after the Ga valence lexicon. While the formal frames could be constructed in parallel and much content could probably be shared, a transfer of the embedded grammatical information (such as syntactic and semantic frames) must be preceded by investigations item by item, and with room for possibilities of non-matches from the item inventory of one language to that of the other. Our comparison of the potential uses of *ba* across the languages – which even carry the same morphological form – gives a flavor of what such a workflow would have to deal with.

²⁷ Lexical Functional Grammar (Bresnan 2001).

²⁸ Cf. <https://universaldependencies.org/>; for the UDG Bambara Treebank project, see <http://cormand.huma-num.fr/guide.html>.

References

- Anyidoho, Akosua, et al. 2006. *Akan Dictionary*. Pilot project. University of Ghana.
- Beermann, Dorothee. 2015. "Data management and analysis for less documented languages." In: Mari C. Jones (ed.) *Language Documentation and New Technology*. Cambridge University Press, 81–94.
- Beermann, Dorothee & Pavel Mihaylov. 2014. "TypeCraft collaborative databasing and resource sharing for Linguists". *Languages Resources and Evaluation* 48, 203–225.
- Beermann, Dorothee, Lars Hellan & Tormod Haugland. 2018. "Convergent development of digital resources for West African Languages". In: Claudia Soria, Laurent Besacier & Laurette Pretorius (eds.). *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.
- Beermann, Dorothee & Lars Hellan. 2018. "West African Serial verb constructions: the case of Akan and Ga". In: Augustine Agwuele & Adams Bodomo (eds.). *The Routledge Handbook of African Linguistics*. London and New York: Routledge, 207–221.
- Bresnan, Joan. 2001. *Lexical Functional Grammar*. Oxford: Blackwell.
- Brobby, Susana. 2015. Codeswitching on Ghanaian Radio Talk-show: "Bilingualism as an Asset". Master's thesis, University of Bergen, Norway.
- Bruland, Tore. 2011. Ga_verb_dictionary_for_digital_processing. https://typecraft.org/tc2wiki/Ga_Valence_Profile.
- Christaller, Johann Gottlieb. 1881 (latest edition of 2013). *Dictionary of the Asante and Fante Language*. Basel: Basel Evangelical Missionary Society.
- Creissels, Denis. 2015. "Valency properties of Mandinka verbs". In: Andrej Malchukov & Bernard Comrie (eds.). *Valency classes in the world's languages: A comparative handbook*, Vol. 1. Berlin: De Gruyter Mouton, 221–259.
- Dakubu, Mary Esther Kropp, 2004a. "The Ga preverb *ke* revisited". In: Mary Esther Kropp Dakubu & Kweku Osam, (eds.). *Studies in the Languages of the Volta Basin* 2. Legon: Linguistics Dept., 113–134.
- Dakubu, Mary Esther Kropp, 2004b. "Ga clauses without syntactic subjects". *Journal of African Languages and Linguistics* 25(1), 1–40.
- Dakubu, Mary Esther Kropp, 2008. "Ga verb features". In: Felix Ameka & Mary Esther Kropp Dakubu (eds.). *Aspect and Modality in Kwa Languages*. Amsterdam & Philadelphia: John Benjamins Publishing Co., 91–134.
- Dakubu, Mary Esther Kropp, 2009. *Ga-English Dictionary with English-Ga Index*. Accra: Black Mask Publishers.
- Dakubu, Mary Esther Kropp. 2010. Ga Toolbox project expanded with Construction Labeling valence information. Ms.
- Dakubu, Mary Esther Kropp. 2011. Ga Verbs and their constructions. Monograph ms, University of Ghana.
- Dakubu, Mary Esther Kropp, Lars Hellan & Dorothee Beermann. 2007. "Verb Sequencing Constraints in Ga: Serial Verb Constructions and the Extended Verb Complex". In: Stefan Müller (ed.). *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*. Stanford: CSLI Publications. (<http://csli-publications.stanford.edu/>)
- Dakubu, Mary Esther Kropp & Lars Hellan. 2016. "Verb Classes and Valency Classes in Ga". Presented at SyWAL II (Symposium on West African Languages), Vienna, 2016.
- Dakubu, Mary Esther Kropp & Lars Hellan. 2017. "A labeling system for valency: linguistic coverage and applications". In: Lars Hellan, Andrej Malchukov & Michela Cennamo (eds.). *Linguistik Aktuell/ Linguistics Today* 237: Contrastive Studies in Verbal Valency. Amsterdam/Philadelphia: John Benjamins, 110–149.
- Hellan, Lars. 2019. "Situations in Grammar". In: James Essegbey, Dalina Kallulli & Adams Bodomo. (eds.). *The grammar of verbs and their arguments: a cross-linguistic perspective*. Studies in African Linguistics. Köln: R. Köppe.
- Hellan, Lars & Mary Esther Kropp Dakubu. 2010. "Identifying Verb Constructions Cross-Linguistically". *Studies in the Languages of the Volta Basin* 6(3). Legon: Linguistics Dept., University of Ghana. http://www.typecraft.org/w/images/d/db/1_Introlabels_SLAVOB-final.pdf.
- Hellan, Lars & Dorothee Beermann. 2014. "Inducing grammars from IGT". In: Zygmunt Vetulani & Joseph Mariani (eds.). *Human Language Technologies as a Challenge for Computer Science and Linguistics* [Series Lecture Notes in Artificial Intelligence, Vol. 8387]. Springer, 538–547.

- Hellan, Lars, Dorothee Beermann, Tore Bruland, Mary Esther Kropp Dakubu & Montserrat Marimon. 2014. "MultiVal – towards a multilingual valence lexicon". In: Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.). *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*.
- Hellan, Lars, Andrej Malchukov & Michela Cennamo (eds.) 2017. *Contrastive Studies in Verbal Valency*. Amsterdam & Philadelphia: John Benjamins.
- Levin, Beth. 1993. *English Verb Classes and Alternations*. Chicago IL: University of Chicago Press.
- Malchukov, Andrej L. & Bernard Comrie (eds.). 2015. *Valency classes in the world's languages*. Berlin: De Gruyter Mouton.
- Osam, Emmanuel Kweku 1994. Aspects of Akan Grammar. A Functional Perspective. Ph.D. thesis, University of Oregon.
- Pollard, Carl & Ivan Sag. 1994. *Head-driven Phrase Structure Grammar*. University of Chicago Press.
- Schaefer, Ronald B. & Francis O. Egbokhare. 2015. "Emai valency classes and their alternations. In: Andrej Malchukov & Bernard Comrie (eds.). *Valency classes in the world's languages*. Berlin: De Gruyter Mouton, 261-298.
- Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*. Paris: Klincksieck.
- Van Dommelen, Wim & Dorothee Beermann (forthcoming). Tonal properties of the Akan particle 'na'. *Proceedings of the 19th International Congress of Phonetic Sciences 2019*.

Our resources are Open Access

TypeCraft Akan corpus, Release 1.0: https://www.researchgate.net/publication/323998547_TypeCraft_Akan_Corpus_Release_1.0

Ga Valence Profile and files:
https://typecraft.org/tc2wiki/Ga_Valence_Profile
 Ga in *MultiVal* – an online multilingual valence base:
http://regdili.hf.ntnu.no:8081/multilanguage_valence_demo/multivalence
 Further TypeCraft Akan corpora, non-curated:
<https://typecraft.org/tc2wiki/Special://TypeCraft/PortalOfLanguages>

Abbreviations and labels

Constructional level labels used in examples in (2):

evSuExp	extended verb complex with experiencer subject	suAg	subject is agent
obPostp	object is postpositional phrase	suEmitter	subject is emitter
obLoc	object is locative	suIDobSpec	subject is identical to object's specifier
obEndpt	object is endpoint	suLocus	subject is locus
ob2Endpt	object 2 is endpoint	suPostp	subject is postpositional phrase
obBPobSpec	object is bodypart of object's specifier	suTh	subject is theme
pv1intr	preverb 1 is intransitive	trComp	transitive with sentential complement
pv1obDrop	preverb 1 has a dropped object		

Abbreviations used in glosses

1,2,3	1 st , 2 nd , 3 rd person	FUT	future
AOR	aoist	HAB	habitual aspect
CONS	consecutive	IPFV	imperfective
DEF	definite marker	INGR	ingressive
FOC	focus	NARR	narrative

NEG	negation	REL	relative
NOM	nominalizer	SBJ	subject
PAST	past	SG	singular
PL	plural	VEN	ventive
POSS	possessive		

POS labels

ADJ	adjective	NUM	numeral
ADV	adverb	PNposs	possessive pronoun
ADVplc	place adverb	PRT	particle
COMP	complementizer	V	verb
DEM	demonstrative	V1	V1 in a serial verb construction (SVC)
DET	determiner	V2	V2 in an SVC
N	noun	V3	V3 in an SVC
Np	proper name	V4	V4 in an SVC

Situation labels used in (2)

EXHIBPROPTY	exhibited property
INCREMRESULT-PROPTY	property resulting from an incremental process
MOTIONDIRECTED	directed motion
PHENOM	phenomenon
REDUCTION	reduction
SENSATION	sensation

Appendix 1. Distribution of verbs over valence frames in the Ga lexicon

Distribution over valence frames	No. of verbs
Verbs that occur in intransitive frame only	51
Verbs that occur in transitive frame only	144
Verbs that occur in ditransitive frame only	4
Verbs that occur in intransitive and transitive frame	44
Verbs that occur in transitive and ditransitive frame	9
Verbs that occur in intransitive, transitive and ditransitive frame	6
Verbs that occur in SVC only	20
Verbs that occur in transitive and SVC frame	15
Verbs that occur in intransitive, transitive and SVC frame	14
Verbs that occur in intransitive, transitive, ditransitive and SVC frame	3

Appendix 2. The most frequently used situation types in verb entry specifications in the Ga lexicon

Situation type	Number of entries
ABSENT	29
ACQUISITION	29
CARETAKING	12
CAUSATIVE	23
CAUSED	17
CLOSING	4
COGNITION	83
COMMUNICATION	178
COMPARISON	29
COMPLETEDMONODEVMNT ²⁹	6
CONTACT	56
CREATION	14
CUTTING	19
EJECTION	15
EMOTION	29
EXPERIENCING	45
MAINTAINPOSITION	25
MOTION	180
MOTIONDIRECTED	55
PHENOMENON	29
PLACEMENT	53
POSTURE	7
PROPERTY	164
DYNAMIC-PROPERTY	13
PSYCHSTATE	23
REMOVAL	47
SENSATION	16
TRANSFER	47
USINGVEHICLE	5

²⁹ For 'completed monotonic development (process)'.

Appendix 3. Non-inflected and inflected forms of Akan *ba* 'come' as single verb, in a VV form, or in an SVC, based on specifications in the TypeCraft Akan annotated corpus

	Word	Morpheme	Glosses	English Glosses
UN-TENSED	ba	ba		come
-subject clitic	yɛba	yɛ+ba	1P.PL	we come
TENSED	aba	a + ba	PRF	have come
	baa	ba + a	PAST	came
-subject clitic	wɔaba	wɔ-a-ba	3PL.PRF	they have come
	wɔbaa	wɔ-ba-a	3PL.PAST	they came
	ɔbɛba	ɔ-bɛ-ba	3SG.FUT	he will come
VV – formation	akɔba	a+kɔ+ba	PRF	go + come
	bepuee	ba*+pue-e	PAST	come + reach
Ventive**	ba			deixis – towards a speaker or a location
Serial-Verb Construction (SVC)	ba			
- as VV in V2 slot	abɛtena	a-bɛ-tena	CONS***	come+sit
- as V2	ba			bring (often in combination with <i>de</i> as V1)

* *ba* has several allomorphs when combined with other verbs; *bɛ* is one of them.

** The *ventive* is a type of verbal deixis using the verb *ba* 'come'. It indicates motion towards a specific person or a specific location.

*** The *CONsecutive* is a prefix on the verb following a V1 which expresses continuous aspect or future tense.

Part III



Syntactic structures and conceptualization patterns

■ On the *shi-* evaluative prefix in Lamnso'

Abstract

This paper examines the evaluative features of the *shi-* nominal prefix in Lamnso'. The paper takes advantage of the literature on *shi-* as the noun class marker for Class 19 nouns in Lamnso' to introduce it as a diminutive prefix with the general semantic interpretation of 'small'/'little' (for the same language). The *shi-* prefix is identified here as a polysemous category with varying semantic values. The paper reveals that the *shi-* diminutive introduces pragmatic features yet unstudied for languages like Lamnso'. When the *shi-* prefix combines with the phonological process of reduplication, the limitative rather than the multiplicative interpretation is derived. *Shi-* combines with adjectives to introduce scales of the properties denoted by the adjective. The paper claims that the study of evaluative diminutives across language groups has the potential of revealing affixed valued features that bring to bear the linguistic significance of size and the pragmatics of evaluation.

Keywords: *shi-*diminutive, evaluative morphology, pluractionality, morpho-semantics, Lamnso',

1. Introduction

Lamnso'¹ is spoken in the greater part of Bui Division, which is 150 kms from Bamenda, the capital of the North West Region of the Republic of Cameroon. It is also spoken in Nigeria, specifically in Sarduana Local Government Area of Taraba State. Lamnso' belongs to the Ring subgroup, which is subsumed under the Grassfield branch of the Southern Bantoid languages of the Niger–Congo language family. Southern Bantoid languages are considered to be Narrow Bantu, whereas Lamnso' and other Grassfield languages are non-Bantu (or Wide Bantu). Kom, Oku, Aghem, Babanki and Noni are other languages of the subgroup (Welmers 1973: 159). Lamnso' and Oku are closely related. Lamnso' does not have prominent dialectal variations (Grebe 1984). Like Fula (Arnott 1970), Swahili (Mkude 1990; Welmers 1973) and many other languages of the Wide Bantu family, Lamnso' nouns

¹ Orthographically, the name of the language has been represented in the literature as Lamnsoq, Lam Nso, Lamnso, Nso, Lamnso and Lamnso?. Gradually, the last option has been preferred by most researchers because the variety of existing spellings stem from the struggle to appropriately represent the glottal stop which appears in every lexical position except as a C₁. For typographical reasons, this sound [ʔ] has come to be represented in the literature by [']. In this paper, therefore, **Lamnsoʔ** is written simply as **Lamnso'**.

and nominals fall under different classes on the basis of agreement operated by concord markers which vary from one class to another (Doke 1943, Grebe & Grebe 1975, Eastman 1980, Yuka 1998).

Lamnso' like many Niger-Congo languages has enclitics on virtually all verbs. These enclitics perform different functions in a sentence. Apart from signifying the morpho-syntactic relations existing between the verbs and the arguments, they indicate subject-verb agreement, case marking as well as other grammatical relations that exist among clausal elements (McGarrity & Botne 2002). Yuka (2000) has shown the various interlocking agreement relations identifiable in basic Lamnso' clauses. Such agreement relations are linked to the nominal noun class markers and the verbal enclitics that are a prominent feature of the language. Yuka (2008) observes that the understanding of Lamnso' grammar largely depends on the understanding of its nominal and verbal system.

It could therefore be problematic for the language learner if a morphologically identical enclitic (*shi-*) surfaces both as a noun class marker and an adjectival marker with distinct semantic and functional roles. The *shi-* prefix is easily identified as the marker of Class 19. We introduce it here as a diminutive prefix and investigate its adjectival and pragmatic functions.

2. The Lamnso' Noun Class System

Lamnso' nouns fall into ten major classes as shown in Chart 1. Following the Bantu system of noun class numbering, it is manifested as follows:

Class	Affix		Gloss
1	ø-	shwà'	'a knife'
2	à-	àshwà'	'knives'
3	-ø	wum	'egg'
5	Ø	luj	'a song'
6	me-	mejijy	'stars'
7	ki-	kikun	'a bed'
8	vi-	vikun	'beds'
9	-ø	yo	'a snaek'
10	-si	yosi	'sneaks'
19	shi-	shijiy	'a star'

Chart 1. Nominal affixes in Lamnso'

Classes 6, 7, 8 and 19 take CV prefixes while classes 1, 3, 5 and 9 have no affix on the noun. Class 2 is marked by a V affix and 10 by a CV suffix. The singular/plural marking reduces the nominal system in (1) into six main genders (1/2, 3/10, 5/10, 7/8, 9/10 and 19/6) as demonstrated below :

(1) a)	shwà/àshwà'	ø -shwà/à-shwà'	(1/2)	'knife/knives'
b)	wum/wumsi	wum- ø /wum- si ²	(3/10)	'egg/eggs'
c)	luŋ/luŋsi	luŋ- ø /luŋ- si	(5/10)	'son/songs'
d)	kikun/vikun	ki -kun/ vi -kun	(7/8)	'bed/beds'
e)	yo/yosi	yo- ø /yo- si	(9/10)	'sneak/sneaks'
f)	shijiy/mejiy	shi -jiy/ me -jiy	(19/6)	'star/stars'

The semantic features shared by the nouns in the same class are not consistent. It should be noticed that the singular affix for nouns in Class 3, 5 and 9 commonly take the class 10 affix (-**si**). This demonstrates that the semantic relationships between the Lamnso' noun classes are simply not definitive. The vowels of the noun class prefixes do not bear tone³. The class 19 prefix (**shi-**) is more relevant to the thrust of our discussion because it is morphologically identical to the diminutive prefix (**shi-**), which has been identified as evaluative of the noun rather than an identity of a noun class (Stump 1993).

2.1. *Shi-* as class 19 noun class marker

Banboye (1997) groups Lamnso' nouns into 4 classes (rather than the 6 classes shown in Chart 1 above) and identifies the **shi-** prefix as marking Class 4 nouns. He also classifies nouns with the **me-** (plural) prefix as Class 4 nouns. More importantly, Banboye's claim is that "... most nouns with the **shi-** (prefix) are names of small things, animals or insects..." His Class 4 nouns therefore constitutes the "diminutive" class, which he contrasts with mass nouns that take the **me-** prefix. These observations about **shi-** are generally reasonable but specifically incorrect. It is obvious that Banboye's nominal classification lumps two

² The hyphen and boldface here are simply for clarity.

³ Lamnso' exhibits eight lexically significant tones. Such contrastive tones vary in pitch which is semantically significant. Grebe & Grebe (1975) and Grebe (1984) have done an extensive study of tone in Lamnso'. The language has three level tones:

(1) High tone [˥] as in **kán** 'monkey' **kún** 'beans'

Mid tone [˨˨] as in **way**2 'market' **loŋ**2 'horn'

Low tone [˩] as in **mbàm** 'money' **làv** 'thread'

As a result of some tonological processes in a sequence of the type HL, LH etc. registered contour tones are derived. The language has five glide tones:

(2) Mid-High [˥˨] as in **wum**2.1 'egg' **tinin**2.1 'cut'

High-Mid [˥˨˨] as in **nyoo**1.2 'vegetable' **wuy**1.2 'hair'

High-Low [˥˩] as in **rim**1.3 'witch' **saŋ**1.3 'dry'

Mid-Low [˨˩] as in **wu**2.3 'he/she' **gon**2.3 'perennial illness'

Mid-Low-High [˨˩˥] as in **yaa**2.3.1 'grandmother' **ven**2.3.1 'you'

The mid tone is not marked in language. The orthographic convention requires that only low and high tones be marked in Lamnso' over the vowels within the syllables that bear them. In this paper we have indicated mid and contour tones on the vowels of the verbs to overtly represent a sequence of tones that we discern as relevant to our discussion. Lamnso' tones fall into eight tone classes as shown in the data above. The contrastive tone is always on the nuclear syllable while the tone on the peripheral syllable is always mid-low. The high tone fluctuates freely with the mid-high while the low tone fluctuates with the mid-low (see Grebe & Grebe (1975), Grebe (1984) and Yuka (2012) for a detailed discussion of tone in Lamnso').

classes into one because he misinterprets ‘gender⁴’ for noun class. Banboye’s (1997:14) examples for his ‘fourth class’ are shown in (2) below.

- (2) a) shisi/mesi cat/cats
 b) shinón/menón bird/birds
 c) shinjaàr/mejaàr smallbird/smallbirds
 d) shingwán/mengwán chigoe/chigoes
 e) shilív/melív heart/hearts
 f) shindzà'/mendzà' yam/yams
 h) shilu'/melu' palmwine

The above examples indicate that *shi-* does not necessarily signal small things, animals or insects as can be discerned from the gloss of *shilív* ‘heart’, *shinón* ‘bird’, ‘small bird’ etc. For instance, there is not a concept for a ‘small heart’ in Lamnso’. Any two-legged winged animal (with a beak) is conceived of as a bird. But there are species of birds. *Shinjaàr* happens to be the Lamnso’ name for weaverbird⁵. Also, *shilu'* ‘palm wine’ is a reference to one container of palm wine (a calabash or jug). *melu'* (palm wine) is a mass noun which takes the Class 6 nominal prefix that happens to be the plural of Class 19 nouns as shown in example (2).

The more interesting observation is that number agreement within the Lamnso’ clause is triggered by the nominal prefix of the argument in subject position (Yuka 1998) as shown in examples 3a-b and 4a-b.

- (3) a) **shi-** nón **she** sánér **shin shi** dzə **shém**
 cl.19 bird cl.19 AM black cl.19 DET-these cl.19 sm cl.2-is cl.19 POSS
 ‘This blackbird is mine’
 b) **me-** nón **me** sánér **men** **me** dzə **meém**
 cl.6 bird cl.6 AM black cl.6 DET-this cl.6sm cl.2-is cl.6 POSS
 ‘These blackbirds are mine’
- (4) a) **shi-** lu' **she** kù'un **shin** **shin** nyòm shaà **shisó**
 cl.19 palmwine AM19 big cl.19 DET-this cl.19 sm cl.2 sweet cl.3 pass 19DET-that
 ‘This big (calabash of) palmwine is sweeter than that (one)’
 b) **me-** lu' **me** kù'un **men** **me** nyòm shaà **mesó**
 cl.6 palmwine AM6 big cl.6 DET-this cl.6 sm cl.2 sweet pass DET-those
 ‘These big (calabashes of) palm wine are sweeter than those (ones)’

Examples (3) and (4) show that each nominal prefix triggers a chain of agreement relations in a Lamnso’ sentence. The morphology of the basic constituents is agreement

⁴ Nominal categories of nouns and/or pronouns are organized in masculine, feminine, neuter, or common sets.

⁵ A gregarious finch common to Africa and Asia, known to live in communal nests. Ploceidae is its scientific name.

dependent. For instance, only a thorough analysis of the internal structure of the Lamnso' Determiner will reveal that its base form is *-n*, just as the subject marker that resumes the nominal in subject position is morphologically identical to the noun class marker (*shi/me*). More importantly, Lamnso' syntax is agreement dependent. The only way to understand lexical, phrasal and clausal computations in this language is to unravel the agreement relationship in the language triggered by the noun class prefixes (Yuka 2000).

3. Conceptual orientation

From a general point of view, evaluation is conceived here as an interpretation of the primitives of qualitative or subjective appraisal of GOOD and BAD. It includes the perception of BIG and SMALL presented in objective/subjective terms or purely from its qualitative/descriptive attributes. An evaluation is a call to judgment of the assessor over an object.

The evaluation can be either descriptive/objective or qualitative/subjective. Evaluative categories include: Profession, Education, Experience etc. The object is also valued from a pragmatic perspective which is dependent on the intention(s), viewpoint and evaluation standard of the evaluator (Dressler & Barbaresi 1994). The evaluation of the object is predicate and context dependent (Default Value, Standards, Traditions, Speech Community). Evaluation is a mental operation that employs relative parameters to value an object. The relativity of evaluation raises the question of the objectivity of an assessment.

Evaluation as a linguistic phenomenon is manifested through the so-called diminutives, augmentatives, pejoratives, etc. (Körtvélyessy 2015; Grandi 2005; Grandi 2009). This paper examines the functional value of the *shi-*diminutive prefix and the concept of evaluation via size in Lamnso'.

4. *Shi-* as an evaluative prefix

The nominal features of *shi-* as the Class 19 noun class marker in Lamnso' have been introduced in section 2 above. Functionally the diminutive *shi-* is akin to Lamnso' verbal extensions that alter the internal structure of the event specified by the verb. As a noun class marker, *shi-* is non-evaluative, it distinguishes a noun with its morphological system of agreement. But *shi-* may be used as a morpheme added to other nouns. Banboye's (1997) observation that prefixing *shi-* to nouns from other noun classes is one of the strategies Lamnso' employs to identify small or diminishing things is (in part) correct. This observation draws our attention to the evaluative values (Fradin et al. 2009; Dany 2014) of the nominal to which *shi-* is prefixed. This section of the paper also seeks to determine the possibility that the evaluative features ascribed to the noun through the *shi-*prefix can be extended beyond the semantic value of size. The *shi-*prefix tends not only to assume adjectival features but can also be interpreted as an adjectival qualifier that introduces attitudinal evaluation extending beyond the linguistic structure of the utterance. This pragmatic evaluation appears to be the principal semantic value of the *shi-*prefix in Lamnso'. The *shi-* prefix in the language appears to have a multiplicity of functions and interpretations.

We examine these functions in section 5 below towards assessing their evaluative features and their functional impact on both the qualifiers and the qualified nominals.

4.1. *Shi-* as a marker of diminutives

Shi- as an evaluation affix introduces the concept of diminution. The prefix *shi-* is affixed to Lamnso' nouns that (from the speaker's view) should be bigger than their current size. In this context, *shi-* is therefore a nominal qualifier and not a noun class marker. In the data below, each noun is first presented in its bare form before the addition of the prefix.

(5)	tu'	'potatoes'	shitu'	'little/small quantity of potatoes'
	ndzǎreŋ	'groundnuts'	shindǎreŋ	'little/small quantity of groundnuts'
	kibam	'a bag'	shibamin	'a small bag'
	kikúm	'a shirt'	shikúmin	'a small shirt'
	kiŋgom	'a banana'	shingomìn	'a small banana'
	meŋkfəŋ	'flour'	shikfəŋ	'a small quantity of flour'
	nyoo	'soup'	shinyoo	'a small quantity of soup'
	m̀bàm	'money'	shimbàm	'a small amount of money'
	mboò	'greater cane rat'	shimboò	'a small greater cane rat'
	ré'	'yam tuber'	shiré'	'a small yam tuber'
	meŋgvár	'oil'	shingvár'	'a small quantity of oil'
	koy	'arm'	shikoy	'a tiny arm'
	wán	'child'	shiwán'	'a little child'

The evaluation of the nouns in (5) is relative to the judgment indices of the speaker (descriptive/objective and qualitative/subjective). For mass nouns (like water, wine, soil etc), the *shi-* prefix is used to refer to single container sizes of such nouns. A few singular nouns with the *ki-* noun class prefix take an additional *-in* suffix when the demunitive *shi-* prefix is added.

4.2. *Shi-* and reduplication

Reduplication is the morphological process through which all or a part of a form is repeated. Generally, repetition has the semantic value of augmentation or emphasis. In Lamnso' this augmentation value of reduplication is altered to diminution once the reduplicated noun takes the *shi-* prefix as shown in (6).

(6)	shitu'	'a little/small quantity of potatoes'	shitutu'	'an very small quantity of potatoes'
	shimboò	'a small greater cane cutter'	shimbòmbò	'a very small greater cane cutter'
	shikoy	'a tiny arm'	shikokoy	'a very tiny arm'
	shimbàm	'a small amount of money'	shibàbàm	'a very small amount of money'
	shingvár	'a small quantity of oil'	shingvǎŋgǎr	'a very small quantity of oil'
	shiwán	'a little child'	shiwáwán	'a very little child'
	shifúm	'a small forehead'	shifúfúm	'a very small forehead'

In example (6) the highlighted stems can be reduplicated. Reduplication introduces emphasis on the diminutive feature of the nominal. The reduplication is partial because only part of the nominal form is repeated given that the final consonant of the noun is deleted in the first part of the reduplicated noun. It therefore turns out that the reduplication highlights the semantic feature of diminution (*shi-*) rather than duplication or the augmentation of the noun it qualifies.

5. The pragmatics of *shi-* as an evaluative prefix

In general discourse, interlocutors often engage in the evaluation of objects of interest. The judgments they make through their utterances tend to reveal a little more than their linguistic structures portrayed. Diminution and augmentation are often used to express endearment or disdain, which are subtle features of evaluation. The motivation(s) for evaluation often cannot be separated from predication and context. The semantic value of a diminutive affix like *shi-* (in Lamnso') cannot be fully understood only within the linguistic structure. This section of the paper subsumes the *shi-* prefix within the relative parameters of evaluation.

Our basic claim about the diminutive prefix *shi-* is that it introduces the feature SMALL to the noun it qualifies. Objects that are diminishing, have diminished quantitatively or in size fittingly take the *shi-* prefix. We have also argued that such evaluation can either be objective or subjective given the intention of the evaluator. In the examples below, we present basic sentences in Lamnso' that portray the *shi-* diminutive prefix and attempt a functional analysis of them.

- (7) *fo mu shibanin i m yiir shiwán shem sho*
 'Give me **little food** to feed my **small child**'

In (7), the Speaker is requesting for *shibanin* (little food) to feed his/her hungry *shiwánn* (small child). In an attempt to persuade the Hearer, the Speaker minimizes the object of his request. He/she isn't necessarily requesting a small quantity of food but enough food to satisfy the hunger of his/her child. The intent of the diminutive qualification of *kiban* (food) as *shibanin* (little food) is persuasive. The speaker does not wish to be seen as asking for too much. The persuasive evaluation of the object eases the possibility of it being acceded to. The child may not be as small as presented by the Speaker. '*shiwán*' here is an endearment, presented in a diminutive vulnerable state that should attract sympathy from the Hearer. The qualification of the food and the child in diminutive terms is intended to facilitate the fulfillment of the Speaker's request.

- (8) *áyii du juu wuna shingəv shi le?*
 'How does one go to pay a bride-price with a **small chicken**?'

In (8), the Speaker questions the rationale for taking *shingəv* (a small chicken) to one's future in-laws. Usually bride-price items are carefully selected to give the impression to your future in-laws that you have the capability to take adequate care of their daughter.

Part of that evidence is in the size and quantity of the items one brings as the bride-price. *Shingəv* in (8) may not necessarily be a small chicken but it is presented in diminutive terms to express dissatisfaction with the expected size of chicken for the purpose the chicken is meant.

- (9) *m dù yen jín vày a dzə shiwán shibor*
 ‘I have seen that bride who is an immature child (**small baby**)’

In (9), the Speaker expresses his/her disapproval of child marriage. The bride is not only qualified as *shiwán* (small baby) but also as immature. The use of *shiwán* in (9) contrasts with the use of the same diminutive in example (7), because unlike in (7), there is no endearment in (9). The attitude of the Speaker in (9) is that of distant disdain to the young bride and the persons endorsing such a marriage. The speaker employs the diminutive to qualify his/her non approval of underage marriage.

- (10) *liù wíy vən bìr fěyì. Moo lón shiŋkfəŋ fo wun wu ŋwey*
 ‘That woman is badly behaved. I begged for a **small (quantity) of corn flour** and she declined to give me’

The utterance in example (10) is judgmental. The Speaker labels the lady referenced in the utterance as a ‘...badly behaved...’ woman simply because the lady failed to accede to her request for *shiŋkfəŋ* (a small quantity of (corn) flour). In reality, the Speaker’s request may not have been as small in quantity as she will have her interlocutor(s) believe. She however presents the size of the object of her request in diminutive terms because her diminutive characterization renders her request easy to be acceded to. The Speaker’s intension is to typify the lady’s failure to comply to an otherwise basic request as bad behavior.

- (11) *à kum shuu ben bì' shítutu' she woo foon mu yoone á?*
 ‘Are you this boastful because of **the very small (quantity) of potatoes** you gave me yesterday?’

In example (11) the Speaker is not happy that his/her interlocutor gave her potatoes and publicized the benevolent act. The quantity of the potatoes in question is presented in diminutive terms (*shítutu'*). The evaluation of the quantity of the object of benevolence may not necessarily be on focus here. The Speaker is rather objecting to the loquaciousness that followed the act of giving. The Speaker’s intention is to register the point that the publicity given to the kind act does not match the sacrifice of the benefactor.

What the utterances in (7-11) indicate is that the semantic evaluation of the *shi-* prefix in Lamnso' is complex. *Shi-* as a nominal prefix cannot simply be pinned to size. The evaluation intentions of the Speaker, his/her world view, the context of the event, the interlocutors and more, all contribute to the correct interpretation of the *shi-* prefix.

6. Conclusion

The Lamnso' *shi*-prefix is clearly a polysemous unit which presents itself on the one hand, as the noun class marker of Class 19 nouns and on the other hand, as an evaluative element of size. The pragmatics of evaluation reveal that the concept of *small/big* becomes fluid when the assessor's pragmatic values of assessment and the default values of the object under evaluation are considered. The claim here is that the *shi*-prefix in Lamnso' connotes more than just the diminutive aspects of the nominal under qualification and very often the prefix portrays the Speaker's attitude in an evaluation of the noun. Most times the emphasis is not about size or quantity, rather the prefix is employed more as a strategy for persuading an interlocutor to appreciate an object from a particular perspective.

References

- Arnott, Daniel. 1970. *The Nominal and Verbal Systems of Fula*. Oxford: Clarendon Press.
- Banboye, William. 1997. *Elements of Lamnso' Grammar*. Bamenda: Copy Printing Technology.
- Dany, Amiotand & Dejan Stosic. 2014. "When evaluative morphology, pluractionality and aspect get tangled up: A case study of French suffix edverbs". In: Zoe Gavriilidou & Anthi Revithiadou (eds.). *Mélanges offerts à Anna Anastassiades-Syméonides à l'occasion de sa retraite*. Kavala: Editions Saita, 16-33. HAL Id: hal-00952491.
- Doke, Daniel. 1943. *Outline Grammar of Bantu*. Mimeographed. Johannesburg: Witwatersand University Press.
- Dressler, Wolfgang & Lavinia Merlini Barbaresi. 1994. *Morphopragmatics. Diminutives and intensifiers in Italian, German, and other languages*. Berlin: Mouton de Gruyter.
- Eastman, Carol. 1980. "Concordial Agreement in Lamnso'". *African Marburgensia* 13, 25-51.
- Fradin, Bernard & Fabio Montermini. 2009. "La morphologie évaluative". In: Bernard Fradin, Françoise Kerleroux & Marc Plénat (eds.). *Aperçus de morphologie du français*. Paris: Presses Universitaires de Vincennes, 231-266.
- Grandi, Nicola. 2005. "Sardinian Evaluative Morphology in Typological Perspective". In: Ignazio Putzu (ed.). *Sardinian in Typological Perspective*. Bochum: Dr. Brockmeyer University Press, 188-209.
- Grandi, Nicola. 2009. "Restrictions on Italian verbal evaluative suffixes: The role of aspect and pluractionality". *York Papers in Linguistics Series* 2, 46-66.
- Grebe, Karl. 1984. The Domain of Noun Tone in Lam Nso. Master's Thesis, University of Calgary, Alberta.
- Grebe, Karl & Winifred Grebe. 1975. "Verb Clusters of Lamnsok". In: Joseph Grimes (ed.). *Network Grammars*. Norman: Summer Institute of Linguistics, 85-105.
- Körtvélyessy, Livia. 2015. *Evaluative Morphology from a Cross-Linguistic Perspective*. Newcastle: Cambridge Scholars' Publishing.
- McGarrity, Laura & Robert Botne. 2002. "Between Agreement and Case Marking in Lamnso'". In: Robert Botne & Rose Vondrasek (eds.). *Explorations in African Linguistics: From Lamnso' to Sesotho*. Bloomington IN: Indiana Linguistic Club, 53-70.
- Mkude, Daniel. 1990. *Towards a Semantic Typology of the Swahili Language*. Tokyo: Institute for the Study of Language and Culture of Asia and Africa, Tokyo University.
- Stump, Gregory T. 1993. "How peculiar is evaluative morphology?". *Journal of Linguistics* 29, 1-36.
- Radford, Andrew. 1988. *Transformational Grammar*. Cambridge: Cambridge University Press.
- Yuka, Constantine. 1998. "The Lamnso' Noun Class System and the Chomskyan Computational 'Machine'". *Research in African Languages and Linguistics* 4(2), 103-130.
- Yuka, Constantine. 2000. Operative Relations in the Lamnso' clause. Ph. D. Thesis, University of Ibadan.
- Yuka, Constantine. 2008. "Verbal Extensions in Lamnso': An Overview". *International Journal of Language Studies* 2(2), 147-172.
- Yuka, Constantine. 2012. "Basic Tense and Aspect Marking in Lamnso'". *Studii de Lingvistică* 2, 251-267.
- Welmers, William. 1973. *African Language Structures*. Berkeley: University of California Press.

Vanessa Chivir Adzer

ORCID 0000-0001-8179-724X

*Department of Languages And Linguistics,
Benue State University, Makurdi*

■ Tiv compounds and compounding

Abstract

Compounding is a common word formation process which functions in many languages of the world. The combination of two or more free words is usually connected with morphological adaptation of single units to the newly created item. The study is aimed at identifying different compounds attested in the Tiv language and explicating the process of their formation. It is revealed that the units they represent are not limited to noun + noun compounds. The language allows compounds to be formed with elements from any word class; thus, producing structures like noun + noun and adjective + noun as well as various phrases involving verbs. Data was collected through interview, elicitation consultation and written documents. The result confirms that compounding is a productive process in Tiv language and the strategies of word compounding are based on both regular and specific rules of word combinations.

Keywords: Compound, compounding, endocentric compounds, exocentric compounds, Tiv

1. Introduction

A grammatical description of language largely operates at two major levels: the linguistic level of morphology, which is the study of word structure and its formation and syntax, which is the study of sentence structure. This paper concentrates on both, as observed in the description and analysis of complex units as well as the interrelation between the phonology, morphology, syntax and semantics.

Words are as important in the study of language as blocks are in the structure of a building. To claim knowledge of a given language requires knowing the words of such a language. Speakers of languages have thousands of words in their vocabulary and continually create new ones to meet their linguistic needs and, in turn, enrich the vocabulary of their languages. They do this by manipulating their linguistic resources to describe new concepts, ideas, objects and situations as the need arises. Thus, on a regular basis, “older forms of language are either modified or replaced by new ones” (Kari 2015: 93).

There are many ways of creating new lexical units in language. Different languages of the world employ different strategies also called word formation. Compounding is one of them, it may involve inflectional or derivational processes, but new words may arise from clipping, which means the formation of a new word by shortening it or some other types of conversion. Some of the processes are more productive than others, depending

on individual languages. This study concentrates on compounding in the Tiv language and is oriented at items which are multiword phrases of different types. They include nominal compounds and nominal phrases, but also verb + noun collocations and even clausal structures. Their common feature is that they are lexicalized phrases which function as lexical units. Various types of compounds have been identified by categorization of their components, but some structural characteristics were also described. Since they are newly created terms, the purpose of this study is documenting the comprehensive but unconscious knowledge that Tiv speakers have of their language, which could be lost if not preserved through description and documentation.

The study adopted the survey method in investigating the facts about Tiv compounds and compounding. Both the primary and secondary sources of data collection were used. Primary data was sourced through interviews and consultation. Unstructured interview questions were used to elicit the needed data from older Tiv speakers in Makurdi, Benue State. Secondary data was sourced from published and unpublished works such as textbooks, journals, theses and internet publications. A total number of one hundred words were gathered and sorted out. The study made use of forty words as a sample representation of the whole.

The study is anchored on the morphological approach to word analysis known as item-and-process propounded in 1933 by Bloomfield, an American Structural Linguist. The approach item and process also referred to as lexeme-based morphology explains the realization of “complex words as resulting from the operation of processes on simpler words” (Aronoff & Fudeman 2011: 49). To further explicate the approach, Agbedo (2015:195) explains, “item-and-process approach, views a word-form as a result of applying rules that alter a word-form or stem in order to produce a new one.” The compounding rule is one of such rules where a word-form is altered to produce a new word. For example: *tea+cup* = ‘teacup’.

2. Tiv people and their language

Tiv is a term that bears three varying meanings. According to Udu (2009: 1), “first, it refers to the ancestral father of Tiv race of Central Nigeria. Second, it refers to the ethnic group which is a descendant of Tiv. Lastly, it refers to the language spoken by the race.” The concern of this study is basically with the last two meanings.

The Tiv people are a unique ethnic group of semi-Bantu linguistic affinity. They speak one common language having varieties of realization but mutually intelligible to all. Hence, the language is more or less homogenous. The Tiv are one amongst many ethnic groups in Nigeria. They are the most populous ethnic group in the Middle Belt area of Benue. Though found in many other places, they predominantly reside in Benue State, with their traditional headquarters in Gboko. As Udu (2009: 1) says

the Tiv people are found in Central Nigeria and a significant number in the Republic of Cameroon in West Africa in the Continent of Africa. The Tiv people occupy over thirty three (33) local government areas across Benue, Nassarawa, Taraba and Cross River States [...]. Tiv is spoken by well over 3 million people.

Tiv people share boundaries with Jukun in the North, Cross River on the East, the Idoma in the South and Lafia in the West. In Benue State, they share their land with other ethnic groups such as, Jukun, Nyifon, Abakwa, Etulo and Iyon. They live in clusters and are neighbourly.

2.1. Aspects of Tiv grammar

Any aspect of the study on Tiv grammar is conditioned by the features of its sound system and typological patterns of its morphology. As far as phonology is concerned, the Tiv language has fifty three sounds, divided into thirty three consonants and twenty vowels (Shoja 2010: 1). Among the vowels, twelve are pure vowels while eight are diphthongs. Tiv operates a consonant-vowel (C-V) structure in word formation.

The morphology of Tiv has been the subject of linguistic studies since the 1940s. In the early descriptions, noun classes have been recognized as a morphologically marked category (Abraham 1940; Welmers 1973). Similarly, verb classes which determine the complex morphology, have been distinguished (Arnott 1958). The complex derivational patterns, however, are still interpreted differently in terms of their status and the function of morphemes. Udu (2009: 41) opines that Tiv morphology is highly inflectional insofar as the inflectional affixes do not change the grammatical class(es) of the stem(s) to which they are attached, but signal grammatical relationships such as plural, change in tense and possession.

2.1.1. Tiv word structure

The formation of words in Tiv follows the rudiments of word formation common to other languages. The structure of the words includes syllables of V, VC, CV and CVC patterns and their various combinations. According to Shoja (2010: 39), the structure of Tiv words ranges from one-syllable words to multi-syllabic words, e.g. *fa* 'know', *oo* 'sunshine', *or* 'man, human-being', *a-to* 'ears', *tahav* 'power', *gbargbar* 'swallow', *kohol* 'gather', *a-wam-be* 'blood', *i-kya-kya-se* 'tremble', *a-mbor-a-vu-ngu* 'myth', *i-kyo-nov-ki-tor-ov* 'royal chair', *u-gbi-lin-mba-i-or-ov* 'useless men'.

In terms of morphological formation, Tiv words can be described as comprising a simple structure (words with a single stem that are indivisible), for example *kon* 'wood', a complex structure (words with one or more morphemes attached to the stem), for example *mba-alum* 'oranges' and also a compound structure (two stems combined to form a word), as in *or-twer* 'doctor'.

Morphemes attached to the stem determine the syntactic function of the word and its grammatical features, for example the prefix *mba-* in *mba-alum* 'oranges' marks plurality (*alum* 'orange'), but other words have different plural morphemes: *i-or* 'persons' (*or* 'person'); *a-kpa* 'bags' (*ikpa* 'bag'). Most nouns require a prefix to mark grammatical number (plurality); a few modify their internal structure to form their plural forms. Some nouns in Tiv can be derived from adjectives by attaching a prefix and a suffix to their stems. The stem *doo* 'beautiful' is an adjective from which the noun *m-doo-m* 'beauty' is derived; and *kehe* 'fat' from which *m-kehe-m* 'fatness' is derived.

In the verbal paradigm, the morphemes attached to the verb stem may have the status of clitic pronouns, as in *na-m* 'give me' (*na* 'give'), *ker-em* 'find me' (*ker* 'find') or category

markers, such as the enclitic, *-n*, which marks the verb for present or past participle (Udu 2009: 41), e.g. *vine-n* ‘dancing’ (*vine* ‘dance’), *ngor-on* ‘shouting’ (*ngor* ‘shout’).

However, among the most characteristic features which also affect the morphology of this language is tonality. Tone is a prominent feature in Tiv. It plays a vital role in distinguishing meaning of some lexical items as well as provides grammatical information like tense and number. Wrong marking of tone or lack of it results in ambiguity, particularly in the written form of the language.

According to Udu (2009: 14), Tiv exhibits three contrastive tone levels:

- (1) Tiv: H (high) – as in *ná* ‘give’
 M (mid) – as in *sēndē* ‘strangle’
 L (low) – as in *ishàn* ‘star’

An obvious example of tone distinguishing meaning among words that have identical forms is:

- (2) Tiv: *wùá* ‘grind’
wúà ‘kill’
wúā ‘sorghum’

Words which are formed from more than one syllable may consist of syllables with different tonal levels. For example:

- (3) Tiv: *á-wám-bē* – ‘blood’ (high, high, mid)
á-kē-sē – ‘baskets’ (high, mid, mid),
ì-jàn-dè – ‘ladder’ (low, low, low).

3. Compounding as a universal phenomenon

Compounding is a word formation process concerned with the combination of lexical items to create new word-forms that function as single units. Compounding occurs when two or more words are joined to make a longer word. It is among several morphological processes by which new words are created and added to the lexicon of the language. The process of compounding uniquely forms new words by manipulating the linguistic resources of the language to create new words with meanings often distinct from its constituent parts.

Yule (2010: 55) defines compounding as “the process of joining two separate words to produce a single form.” Hamawand (2011: 11) further explains,

Compounding is also called composition, it is the process of forming a complex structure by combining two or more free morphemes, of same or different word classes. The resultant form which serves to convey a new message is called a compound [...] a structure made up of more than one free morpheme.

In the same vein, McGregor (2009: 93) expatiates that the process of compounding is a situation whereby “two separate words are sometimes joined together to form a single

word, a new word with a new meaning of its own, a meaning that is not entirely predictable from the component words.”

In view of the above, the compounding process simply joins words together from the same or different word classes to form new words with new meanings that may or may not be predictable from their constituent parts. Such new word-forms that are the result of this process are known as compounds.

3.1. Typology of compounds

In linguistics, a compound is a word that consists of more than one unit. It is formed by pairing a string of free morphemes, which are independent words by themselves and may emanate from the same or different word classes, to form a single word. Such new word-forms belong to a particular word class. The meaning of the compound may be similar to or different from the meanings of its components in isolation.

A compound expresses a single thought. It functions as a single word and it occupies a single grammatical slot. Compounds can be formed from any word class but commonly with open class words. They are “typically affixed just like any other word, providing evidence that we do store them in our mental lexicon as single words. They are actually formed when two words behave as a single word” (Denham & Lobeck 2010: 194-195).

Compounds are not limited to two words. Syal & Jindal (2008: 92) affirm, “Compounds are formed by joining two or more bases. In some cases the hyphen is used to separate the bases while in other cases, the hyphen appears to have disappeared with the passage of time.” There is no hard rule about whether or not a compound should be hyphenated as long as the compound is considered a single unit.

Compounds can be grouped under different group types. Kari (2015: 100) identifies three major grounds to situate the groups: morphological, syntactic and semantic grounds. According to him, morphological criterion allows grouping compounds into affix and non-affix compounds. Affix compounds are composed with an affix as part of the attachment. Haspelmath (2002: 88) explains, affix compounds “consist of more than one stem plus an affix.” In other words, this is a compound with an affix attached to one of its members. Non-affix compounds are compounds that do not have any affix attached to their forms, they are simply made up of two or more free morphemes from the same or different word classes.

At the syntactic level, compounds are grouped into endocentric and exocentric compounds. Endocentric compounds are those compounds that have obligatory heads. The heads bear the bulk of semantic content and determine the grammatical category to which the new word-form will be assigned. Their constituent parts are modifiers of the head whose presence serve to narrow the meaning of the head, hence, it presents a kind or type of the head. “The meaning of the whole can be inferred from the meaning of their parts” (O’Grady et al. 2011: 130-131).

Endocentric compounds exhibit the modifier and head relationship which can be differentiated with reference to what Kari (2015: 103) calls head-final and head-initial languages. According to this view, a compound can be directional in two senses: one sense involves the position of the head whether on the right or on the left. The other sense involves the direction of the relation between the parts of the compound. In the head-final languages,

the head is at the right position and the left member modifies it. While in the head-initial languages, the head is at the left position and the right member modifies it. An example of head initial positioning can be cited from Idoma and Igede languages of Benue State. English expresses head final positioning while Tiv expresses both head initial and final. The following examples illustrate the systemic differences between Idoma and Igede Noun-Noun compounds and English Noun-Noun compounds:

- (4) Idoma: *Inú- echi* *Inúechi* ‘hospital’
 ↓ ↓
 house medicine
- (5) Igede: *Ubé – upwu* *Ub’upwu* ‘school’
 ↓ ↓
 house book
- (6) English: Lunch-box ‘lunchbox’

The Idoma example in (4) shows that *inú* ‘house’ is the head of the compound while *echi*-medicine is its modifier. Similarly, head initial positioning is expressed in Igede. The example in (5) shows *ube*-house, the left constituent of the compound as the head with *upwu* – book on the right as its modifier. The example in (6) shows that English expresses head final positioning. The head in (6) is ‘box’ which while ‘lunch’ is a modifier of the head. It is worth mentioning that in Igede the compounding involves phonological reduction which affects one of the neighboring vowels in a newly created word.

In exocentric compounds none of the item modifies the other and the meaning of the compound cannot be inferred from the meanings of its constituent parts.

Semantically, compounds are grouped into semantic and non-semantic compounds. Semantic compounds have meanings that can be deduced from their constituent parts while non-semantic compounds are compounds whose meanings are opaque, they cannot be derived from the meanings of their constituent parts. This group has similar features with exocentric compounds.

4. The patterns of Tiv compounds

Compounding is one among the most productive methods in creating new lexical item in the Tiv language. As discussed earlier, the compounding process involves joining together the two or more free morphemes to function as one unit.

Compounding as a word formation process in Tiv is so common that the conventions of using the hyphen when compounds are newly created and dropping it when the compounds have gained currency as practiced in English is not observed in the language. Words are easily and freely compounded following the rules of the language to name things as well as persons when the need arises.

In Tiv, a variety of combinatorial possibilities in the creation of compounds exist and they are not limited to Noun+Noun combinations. Compounding also covers Verb+Verb

combinations and all possible variants of their extensions as seen in 4.2. Peculiar to Tiv is the fact that, even when the constituent parts of the compound vary across word classes and the head word emanate from a different word class other than the noun class, the resultant compound is still a noun. In other words, most compound words in Tiv are noun compounds. Possible combinations are as follows:

4.1. Noun-Noun compounds

The words in (7-9) are composed with the root morphemes *Átè* and *ijìr*, *Iyòu* and *adùá* and *Tòr* and *kwásè* respectively. Each constituent part of the compound belongs to the noun class and the resultant compounds are also nouns.

- (7) Tiv: *Átè* – *ijìr* *Átèjìr* ‘law court’
 ↓ ↓
 hut judgement
- (8) Tiv: *Iyòu-* *Adùá* *Iyòuādùá* ‘church’
 ↓ ↓
 house worship
- (9) Tiv: *Tòr* – *kwásè* *Tòrkwásè* ‘queen’
 ↓ ↓
 King-woman

A syntactic variant of nominal compounds is Noun+Verbal base for Progressive¹ or Noun Adjective combination, the words in (10-12) are formed with nouns and verbal forms used in Progressive, *moron*, *òròn* and *nyòòn* respectively. Although the second members of the compounds, which are the word’s heads, are of verbal origin, thus expressing head final positioning, the resultant compounds are still nouns.

- (10) Tiv: *Iyóu* – *mòrón* *Iyomoron* ‘pride’
 ↓ ↓
 Body swelling
- (11) Tiv: *Kwàgh* – *òròn* *Kwàghòròn* ‘speech’
 ↓ ↓
 Thing speaking
- (12) Tiv: *Ishímá* – *nyòòn* *Ishímányòòn* ‘anger’
 ↓ ↓
 Heart paining

¹ In the traditional descriptions of Tiv verbs (e.g. Arnott 1958: 119) the form with the suffix –n attached to the verb (in 2.1.1. mentioned as a form marked for present or past participle) is recognized as a verbal base to distinguish Continuous/Progressive and certain Habitual tenses from the Past tense.

In (13), the adjectival phrase *zégē mngēr* is a term for ‘sea’, in which *zege* ‘large’ qualifies *mngerem* ‘water’ expressing a large body of water i.e.

- (13) Tiv: *Zégē – mngēr* *Zégēmngēr* ‘sea’
 ↓ ↓
 Large water

4.2. Verbal compounds (Verb+Verbal base for Progressive)

The combination of auxiliary verb and the main verb in Progressive still forms noun compounds as in (14-16). The combination is phrasal in nature, auxiliary verbs having a null subject are joined with nominal form of the verbs. The words in this group are restricted to personal names.

- (14) Tiv: *Ngù – nán* *Ngùnán* (lit. is providing)
 ↓ ↓
 Is – giving

- (15) Tiv: *Mbá – sén* *Mbásén* (lit. are mocking)
 ↓ ↓
 Are laughing

- (16) Tiv: *Ngù –ván* *Ngùván* (lit. is coming)
 ↓ ↓
 Is coming

4.3. Verb – Noun compounds

The combination in (17-18) is between the verb and noun grammatical classes. The right member of the compound is the head of the structure while the left member functions as a modifier of the head. Thus, the verbs *gbēr* and *kēēr* function to modify the nouns *kón* and *ùsù* in the below combinations. Both examples are personal names.

- (17) Tiv: *Gbér – kón* *Gbērkón* ‘cut wood’
 ↓ ↓
 Cut wood

- (18) Tiv: *Kēēr – ùsú* *Kēērùsú* ‘set fire’
 ↓ ↓
 Make fire

It is to mention that in (17) the first root *gbér* carries a high tone when in isolation but changes to mid when combined with the root *kón* to form the compound *gbērkón* ‘cut wood’.

the plural property of the head *ákāā* has become the property of the whole. Similarly, the plural property of the head *mkúngù* which is *m-* compels its modifier *túlé* to also receive the plural feature *u-* with a plural connective *mbá-*.

Therefore, stems with their inflected parts are combined without the necessity to drop their inflected part as seen in examples (22-24).

- (22) Tiv: *Ìshébēē-kwāgh* *Ìshébēēkwāgh* ‘word’
 ↓ ↓
 piece thing
- (23) Tiv: *Á-sémbēē-á-kāā* *Ásémbēēákāā* ‘words’
 ↓ ↓ ↓ ↓
 Pl pieces Pl things
- (24) Tiv: *Ù-túlé-mbá-m-kúngù* *Ùtúlémbámkúngù* ‘young pigeons’
 ↓ ↓ ↓ ↓ ↓
 Pl young Pl Pl pigeon

However, exceptions exist (but are optional) where the inflection for number is deleted in the course of forming a compound. This is seen in (7), where the *i* in the words *ijìr* is deleted to form the compound *átèjìr*. Similarly in (25) when *ityò* becomes a part of *péndátyò*, e.g.:

- (25) Tiv: *Pēndà- ityò* *Pēndátyó* ‘favourite wife’
 ↓ ↓
 Place head

Most of the compounds do not have any affix attached to their forms, they are simply made up of two or more free morphemes from the same or different word classes as shown by the combination. This class of compounds is more numerous in the language than any other class. They are all noun compounds.

As for tonal structure, the constituents maintain their specific tone markings in compounds. Exceptions were presented with the example (17) where the first root *gbér* carries a high tone when in isolation but changes to mid when combined with the root *kón* to form the compound *gbèrkón* ‘cut wood’). The patterns of compounding in which the tonal structure is a sequence of word constituents are presented in (26- 29).

- (26) Tiv: *Ìmbôr – mngēr* *Ìmbôrmngēr* ‘spring of water’
 ↓ ↓
 Spring water
- (27) Tiv: *Sév – áv* *Séváv* ‘day break’
 ↓ ↓
 Try thing

- (28) Tiv: *Kwásè – yá* *Kwásèyá* ‘housewife’
 ↓ ↓
 Woman house
- (29) Tiv: *Òr – ihòm* *Òrihòm* ‘enemy’
 ↓ ↓
 Man hatred

5.2. Syntactic classification

Syntactically, Tiv compounds are classified into endocentric and exocentric compounds.

5.2.1. Endocentric compounds

Tiv endocentric compounds exhibit head-final and head-initial positioning as explained in (3.1.) above. In other words, both the modifier and head of a Tiv endocentric can take either left or right positions in the structure. For example: the compounds in (30-31) have the left constituent part modifying the right, which is the head and which also contains the basic meaning of the whole compound while the modifier part restricts the meaning. The part of speech to which the head belongs forms the part of speech of the whole compound. The adjective *Gógò* modifies the noun *iyòu* and the adjective *Íniun* (which can be a noun in isolation) modifies *kwásè* expressing a type of house and kind of woman respectively.

- (30) Tiv: *Gógò – iyòu* *Gógòiyòu* ‘tall or high building’
 ↓ ↓
 Tall house
- (31) Tiv: *Íniun – kwásè* *Íniunkwásè* ‘virgin’
 ↓ ↓
 Virgin girl

In the following examples (32-33) the heads *Òr* and *kwāgh* nouns are the left members of the whole while the modifiers *tésè* and *témán* verbs are the right members of the whole. They also express the type of ‘man’ and the type of ‘seat’.

- (32) Tiv: *Òr – tésè* *Òrtésè* ‘teacher’
 ↓ ↓
 Man teach(ing)
- (33) Tiv: *Kwāgh – témán* *Kwāghtémán* ‘seat’
 ↓ ↓
 Thing sitting

5.2.2. Exocentric compounds

The compound earlier mentioned under (25) has no head and no modification from any side. Its meaning cannot be determined from its members. The verb *Pēndà* is not

functioning to modify the noun *ityó* nor is *ityó* posing as a head. The combination of these two yields a word with a meaning far from their individual meanings.

- (25) Tiv: *Pēndà-ityó* *Pēndàtyó* ‘favourite wife’
 ↓ ↓
 Place head

Another type of exocentric compound is presented in (34) in which the noun *shímá* is combined with an adjective *ndôhôr* as a bearer of the specified attributes providing the compound. In this connection, the compound *Ndôhôrshímá* has a meaning that cannot be easily deduced from the meanings of the adjective *ndôhôr* and the noun *shímá* respectively, i.e.:

- (34) Tiv *Ndôhôr-shímá* *Ndôhôrshímá* ‘disgusting’
 ↓ ↓
 Cold heart

5.3. Semantic classification

Semantically, Tiv compounds are grouped into semantic and non-semantic compounds.

5.3.1. Semantic compounds

The meanings of the compounds in (35-38) can be deduced from the meanings of their parts. They are a combination of different word classes and the positioning of head and modifier is flexible. Those in (35-36) take head initial position while those in (37-38) take head final position.

- (35) Tiv: *Kwāgh – nyōōn* *Kwāghnyōōn* ‘sickness’
 ↓ ↓
 Thing paining
- (36) Tiv: *Òr – hémēn* *Òrhémēn* ‘leader’
 ↓ ↓
 Man leading
- (37) Tiv: *Gbilin – kwāgh* *Gbilinkwāgh* ‘vanity’
 ↓ ↓
 Useless thing
- (38) Tiv: *Nyían – òr* *Nyíanòr* ‘fair man’(complexion)
 ↓ ↓
 Red man

5.3.2. Non-semantic compounds

This compound type is connected with metaphoric use of one or more parts of the compound, as in example (39). Without this condition, the meaning expressed by *Shímá* – ‘heart’ and *nyíán* – ‘red’ has little relation to the meaning of the whole *shímányián* – ‘anxiety’.

- (39) Tiv: *Shímá-nyíán* *Shímányián* ‘anxiety’
 ↓ ↓
 Heart red

Similarly, the meaning of clausal compounds may refer to the contextual understanding of the whole phrase rather than be derived from the meaning of the constituents, e.g.:

- (40) Tiv: *Òr-nùmún-gbéndá* *Òrnumúngbéndá* ‘robber’
 ↓ ↓ ↓
 Man scatter road

There are other types of Tiv compounds that combine more than two words, they show that the structure of Tiv compounds is not limited to only two words. The compounds in (41-43), despite their three word combination and the placement of the primary stress remaining on the leftward member of the compound and the resultant words, are all noun compounds.

- (41) Tiv: *Òr – kwásè – hé* *Òrkwásèhé* ‘groom’
 ↓ ↓ ↓
 Man woman new

- (42) Tiv: *Ìshímá-véréń-kéghén* *Ìshímávéréńkéghén* ‘hope/expectation’
 ↓ ↓ ↓
 Heart keep waiting

- (43) Tiv: *Ímbúsú-tár-iyán* *Ímbúsútáryán* ‘east’
 ↓ ↓ ↓
 Side world right

Creating new words on the basis of existing items functions in many languages. The examples from Tiv present it as a combination of two or more free words to function as single word, but the rules which they follow refer to both regular and irregular word combinations. The meaning of these compounds is established by the users. The compounds which have a status of nouns are created from the constituents representing different parts of speech.

According to Umaru (2016:19), similar combinations and categorisations exist in Idoma language of Benue state Nigeria. For instance: the example in (44) shows the case of endocentric compound where the meaning of the whole is simply the sum of the parts. While (9b) is an example of exocentric compound where the meaning of the whole is opaque and cannot be deduced from its parts.

- (44) Idoma: *Ehó – hòì* *ehohoi* ‘goodnews’
 ↓ ↓
 Message good
- (45) Idoma: *Okpá – akpà* *okpakpa* ‘righteousness’
 ↓ ↓
 Stream wood

6. Findings and concluding remarks

The study investigated compounding as a viable word formation process in the Tiv language. The identified compounds were analysed in terms of their features and characteristics distinguishing them from other structures in the language. The study established that:

1. The formation of compounds in Tiv language is the process of joining independent words in one lexical unit, which involves interaction between phonology, lexicon, morphology and syntax. As a tonal language, Tiv also marks the status of newly created items by tonal pattern.
2. Most of Tiv compounds express head and modifier relationship. However, they exhibit both head initial and head final positioning, making the position of head and modifier flexible.
3. The common compound type in the language is a noun compound. It is often the resultant word form regardless of word class combination and the class to which the head belongs.
4. The most productive category of compounds in Tiv is endocentric compounds with meanings that can be deduced from their constituent parts.
5. In most cases of word compounding discussed here, the components preserve markers of their grammatical categories (number, aspect, etc.). That is, stems with their inflected parts are combined without the necessity to drop their inflected part. However, exceptions exist (but are optional) where the inflection for number is deleted in the course of forming a compound.
6. The words also maintain their tone markings even when combined with other roots to form compounds. However, few exceptions exist where the tone of a word may change when combined with another word to form a compound.
7. Compounding is one of the most productive word formation processes in Tiv.

References

- Abraham, Roy Clive. 1940. *The Principles of Tiv*. London.
- Adzer, Vanessa Chivir. 2016. A study of Tiv morphosyntactic relations. Unpublished Thesis, Department of Languages and Linguistics, Benue State University, Makurdi.
- Agbedo, Chris U. 2015. *General linguistics: Historical and contemporary perspective*. Nsukka: KUM-CEE-Ntaeshe Press.
- Amott, David Whitehorn. 1958. “The Classification of Verbs in Tiv”, *Bulletin of the School of Oriental and African Studies* 21(1/3), 111-133.

- Aronoff, Mark & Kirsten Fudeman. 2011 [2nd ed.]. *What is morphology?*. West Sussex, UK: Blackwell Publishing.
- Denham, Kristin & Anne Lobeck. 2010. *Linguistics for everyone: An Introduction*. Boston: Cengage.
- Hamawand, Zeki. 2011. *Morphology in English: Word formation in cognitive grammar*. London: Continuum.
- Haspelmath, Martin & Andrea D. Sims. 2002. *Understanding Morphology*. London: Routledge.
- Kari, Ethelbert Emmanuel. 2015. *Morphology: An introduction to the study of word structure*. Port Harcourt: University of Port Harcourt Press.
- McGregor, William B. 2009. *Linguistics: An Introduction*. New York: Continuum.
- O'Grady, William, John Archibald & Francis Katamba. 2011. *Contemporary linguistics: An introduction*. England: Pearson.
- Yule, George. 2010. *The study of language*. Cambridge: Cambridge University Press.
- Shoja, Ayem. 2010. *Tiv Language in practice: A descriptive approach*. Makurdi: Gold Ink Company.
- Syal, Pushpinder & Dharam Vir Jindal. 2008 [2nd ed.]. *Introduction to linguistics: language, grammar and semantics*. Delhi: PHI learning.
- Udu, Titus Trevor Terver. 2009. *Tiv language: A reference book*. Kaduna: Labari Publishers.
- Umaru, Jacob Yakubu. 2016. The syntax of relative clause in Idoma: A government and binding approach. Unpublished PhD Thesis, Department of Languages and Linguistics, Nasarawa State University, Keffi, Nigeria.
- Welmers, William E. 1973. *African Language Structures*. Berkeley and Los Angeles: University of California Press.

Nigerian Pidgin English phraseology in the context of areal influences

Abstract

The development of Nigerian Pidgin English (NPE) and the expansion of its communicative functions is connected to the use of multiword expressions as new means of expression. The question is to what extent the NPE phraseology is derived from the original lexifier, i.e. English, and what the influence of substratum languages belonging to different genetic and typological groups is. The article presents various types of multiword expressions which are based on English words, but in terms of structural features and conceptualization patterns represent collocations copied from other West African native languages. In this connection, special attention is given to the use of body-related terms in NPE phraseology.

Keywords: Nigerian Pidgin English, phraseology, West African languages, areal features

1. Introduction

Nigerian Pidgin English (hereinafter presented also by acronym NPE) is an English-based pidgin spoken as a *lingua franca* across Nigeria. Being the result of contacts between Europeans and native inhabitants in the coastal area of the Niger delta around 19th century, in contemporary times it became an important medium of communication with its social and regional variants (cf. Caron in this volume).

Linguistic studies on this language are still not much advanced, although the descriptive grammars (Faraclas 1996; Elugbe & Omamor 2007) are now available and characteristics of the language structure have been established (Faraclas 2013a; 2013b). One of the most interesting aspects of linguistic studies on Nigerian Pidgin English is contributing to the knowledge on language development as universal process, but equally important are contact phenomena that occur between the lexifier, i.e. English language and languages of the substratum area in which the language has been developing.

The testimony for growing importance of the language and its communicative role can be found in phraseology. This layer has its significance in the language structure, as it allows to create new units in lexical inventory. As more-in-depth research (Frąckiewicz 2019) indicates, multiword expressions in NPE are very specific units which first and foremost are copied from other West African native languages. Even though they are based on English words, they are newly created collocations which are unique and not consistent either with the rules of English grammar or the patterns of English phraseology. Sometimes they even seem to be created against those rules.

All data used for the analyses was firstly classified in the light of their structural properties, secondly in terms of source domain of their linguistic components, which the particular expression was based on¹. The most common structures are verbal phrases and nominal compounds in which some lexical elements are viewed as a source of calques from native languages.

The interpretation of the data is based on their comparison with similar structures of West African languages, Hausa and Yoruba in particular. These similarities also furnish a basis for claims that Nigerian Pidgin English phraseology represents West African conceptualisation rather than follows English multiword structures or universal creole-like patterns.

In Nigerian Pidgin English some notions are expressed by analytical structures, which are perceived as metaphoric and/or metonymic expressions. They make grounds for expressing grammatical categories, such as gender, but they are also source of many other conceptual models. The use of body part terms in metaphoric functions is the most characteristic feature of NPE phraseology and is given more attention on the following pages.

The selection of examples was based on the meaning of the phrase components which are different than in English. In other words, they constitute a group of lexemes which have an idiomatic character, but at the same time they are not a copy of English idioms. The analysis primarily refers to the sequence of source lexemes rather than the grammatical form of their combination effect. Since the classification of phrasal units is based on categorisation of their constituents, some questions related to this issue will be discussed first.

2. Syntactic categorisation of lexemes in NPE

Categorisation of lexemes in Nigerian Pidgin English is highly problematic. According to Mensah, this is not only a problem of this language, but it is in this language that it is clearly manifested. “The difficulties in classifying the nature of the word are largely due to the fact that the term ‘word’ is used in a variety of senses, which usually are not clearly distinguished” (Mensah 2011: 211). For example, according to Faraclas, the category of adjective is absent in the language and there is no possibility to “distinguish the category of ‘adjective’ from the categories of ‘noun’ or ‘verb’ ” (Faraclas 1996: 128) or simply adjectival meanings are expressed by stative verbs (Faraclas 1996: 30).

As a result of multifunctionality, the lexical items may “belong to more than one grammatical class or part of speech” (Mensah 2012: 171). Moreover, due to the lack of morphological elements in marking the differences between verbs and nouns, it is necessary to rely on syntactic criteria (i.e. the function in the sentence) in their identification. As the same lexeme may represent many grammatical categories, i.e. *chop* ‘to eat’ or ‘food’, for the purposes of the present analysis, some reference points were established in order to define the particular grammatical categories. They are as follows:

Verb – a word used to describe an action, state, or occurrence, which forms the main part of the predicate of a sentence, [...]².

¹ The examples were extracted from various written texts in NPE, as well as dictionaries and linguistic articles indicated in references.

² <https://en.oxforddictionaries.com/definition/verb>

Noun – a word which denotes “concrete entities, whose basic role in syntax is in phrases representing arguments of a verb” (Matthews 2014: 269).

Adjective – stated as non-existing in NPE (Faraclas 1996: 30), the category of adjective is perceived as “any form which is reflected by an English adjective in translation, without reference to its derivation or grammatical function in the language being described” (Welmers 1974: 249).

In addition, in examples extracted from the existing works on NPE, the categories follow the classificatory scheme adopted by the author.

3. Structural patterns of idiomatic expressions

From a structural point of view, the most common idiomatic expressions constitute a sequence of the following constituents: *verb-verb*, *noun-noun*, *verb-noun*. Subtypes involving some particular words can be distinguished. A separate group is the *adjective-noun* pattern which is based on English-like adjectival phrases. On the other hand, the compounds which consist of *verb-adverb*, *noun-adverb*, *adverb-verb*, *adverb-noun*, *adverb-adjective*, *adjective-verb* and *adjective-adverb* were not found among the analysed expressions.

3.1. Verbal compounds

Verbal compounds are the result of a combination of two verbs (Mensah 2011). The analysed sources provide the examples as follows:

Compound	Literal meaning	Translation
<i>born throwaway</i>	born, throw away	not in touch with one’s cultural heritage
<i>chop i chop</i>	eat and eat	mutual corruption
<i>cut and sew</i>	cut and sew	Tailor
<i>die wake</i>	die, wake	recover from a desperate financial situation
<i>dry die</i>	dry, die	HIV/AIDS
<i>du kwanagida</i> ³	do, sleep over	sleep over
<i>krai dai</i>	cry, die	Lament
<i>kwat nkpe</i>	scratch, pay	a prostitute
<i>sidonlook</i>	sit down, look	(be) passive

Table 1. A list of compounds which consist of two verbs.

Expressions distinguished in this group remain multiverb constructions in the sentence structure, but the lexical units of the combination are unrecognizable. They were lexicalised to collocations and gained new meanings which are not related to the original one. As shown in the examples, the two verbs make up the compound of two independent units.

³ The noun *kwanagida* is a lexicalised phrase borrowed from Hausa, i.e. *kwana gida* ‘spend a night at home’.

In some cases the conjunction may occur, but it does not change the status of verbs, which remains the same in the two items. It is worthwhile indicating that newly created phrases (compounds) have the status of either verbs or nouns. This strategy of creating nouns on the basis of verb compounds occurs in Akan, e.g.: *nsɔ hwɛ* (lit. try, see) ‘temptation’. Other examples which involve finite verb forms are presented in Table 6.

3.1. Nominal compounds

The next group of examples includes the combination of lexemes which belong to the category of nouns:

Compound	literal meaning	Translation
<i>babi pancake</i>	babe pancake	‘girl fond of make-up’
<i>bathroom slippers</i>	bathroom slippers	‘rubber flip flops’
<i>bush man</i>	bush man	‘unsophisticated man’
<i>chop money</i>	food money	‘housekeeping allowance’
<i>chop-remain</i>	food remains	‘leftovers of meal’
<i>country paper</i>	country paper	‘passport’
<i>doll baby</i>	doll baby	‘dumb person’
<i>koko waves</i>	koko waves	‘unbrushed hair’
<i>papa battalion</i>	father battalion	‘a man with many children’
<i>shoe maker</i>	shoe maker	‘shoe repairer’
<i>short knicker</i>	shorts knicker	‘shorts’
<i>stone ground</i>	stone ground	‘fall heavily’
<i>sun shade</i>	sun shade	‘sunglasses’
<i>tear race</i>	tear race	‘to run’
<i>tori man</i>	story man	‘journalist’
<i>weather man</i>	weather man	‘poor man’

Table 2. List of compounds which consist of two nouns.

As presented in the above examples, the constituents, in contrast to verbal compounds, are not always syntactically equal, in some cases they have a head-modifier structure in the nominal phrase. The examples manifest the variety of conceptualisation patterns used to create new expressions. In most cases this type of compounding results in the creation of nouns which are names of professions, i.e. *shoe maker*; *tori man*. This pattern also functions in Standard English, i.e. *policeman*, *camera-man* and *candlestick maker*, *barrel maker*, *beer maker*. Some combinations seem to be inspired by Standard English lexicalised phrases, i.e. *sun shade* ‘sunglasses’, which is similar to the lexeme *raincoat* in a functional relation between the two elements (‘protect from’) and has the meaning ‘shade protecting from sun’. In examples *tear race*, *country paper*, *papa battalion* the first element functions as the possessor, while the second is the possessed one, which in Standard English would

rather appear as the following: *race's tear*, *country's paper*, *papa's battalion*. In the next two examples, even though they are structurally similar, the constituents have inverse functions and regarding the result they are based on another Standard English pattern. In the expressions *pepper eye* and *tear eye* it is rather the second element which denotes the possessor, while the first one is possessed and the structures *pepper of eye* and *tear of eye*, respectively, would be their equivalents. Therefore, the genitive constructions in NPE are not marked by morphemes, but by the context. Structurally, these types of constructions with no genitive linker are common patterns of many African languages, including Igbo and Yoruba (but not Hausa). However, the expressions which consist of two nouns are also found in the Hausa language, but they are compounds of syntactically independent constituents, e.g. *bindigā-dādī* (lit. gun-pleasantness) 'ready' (Pawlak 2010: 195), *uwā-ùbā* (lit. mother-father) 'an umbrella organization'.

The questions related to the sequence of nominal components and their functional interpretation is more complex, when particular examples are taken into consideration. The two expressions: *doll baby* and *short knicker* are examples of free combinations of Standard English lexical units which form new phraseological units. While *doll baby* constitutes the reversed colloquial Standard English term *baby doll*⁴, the second expression *short knicker* is a kind of redundant repetition of the two terms for 'shorts'.

Among expressions consisting of two adjectives there is only one: *doti doti* (lit. dirty dirty) 'rubbish'. As a result of reduplication, a new noun was created. It follows the pattern based on reduplication which is a derivational strategy commonly found in West African languages (Frąckiewicz 2019: 85f).

3.1.1. Nominal compounds for marking gender identity

In the group of nominal compounds a small subgroup which has the status of a regular grammatical pattern might be distinguished. These are phrases with the lexemes *man/woman* which are used to mark gender identity, i.e. *man pikin* (lit. man child) 'male child', *wuman pikin* (lit. woman child) 'female child', *man sela* (lit. man seller) 'male seller', *wuman sela* (lit. woman seller) 'female seller'. Analogically, the lexemes *boi* 'boy' and *gel* 'girl' may be used to mark gender, i.e. *house boi* (lit. house boy) 'male servant', *house gel* (lit. house girl) 'female servant', *baby-boi* (lit. baby boy) 'male infant', *baby-gel* (lit. baby girl) 'female infant'.

A nominal compound with lexemes *man* 'man'/'*wuman* 'woman' has a special status in Nigerian Pidgin English phraseology and grammar. Developed on the basis of Standard English, the pattern manifested in *policeman/policewoman* was extended to the gender marker. However, as demonstrated in the examples, the pattern is not directly copied from Standard English. While in examples with *boi/gel* the lexemes follow the noun, in the expressions which consist of *man/wuman* they precede the noun, which is incompatible with the Standard English pattern.

⁴ **Baby doll:** 1. A doll that looks like baby, 2. A short, frilly ultra-feminine dress (<https://www.collinsdictionary.com/dictionary/english/baby-doll>)

3.3. Verb-noun combinations

This group of expressions is the largest one in the corpus of data. The combination of a verb with a noun seems to be the most productive model for word compounding leading to the formation of new lexical items. One of the most productive groups of verb-noun combinations are those based on the verb *chop* ‘to eat’, i.e.:

Expression	Literal meaning	Translation
<i>chop beans</i>	eat beans	fail at something
<i>chop bullet</i>	eat bullet	get shot
<i>chop burger</i>	eat burger	put on weight
<i>chop cockroach</i>	eat cockroach	be pregnant
<i>chop knuckle</i>	eat knuckle	high five
<i>chop life</i>	eat life	to enjoy
<i>chop maut</i>	eat mouth	to kiss
<i>chop slap</i>	eat slap	be slapped

Table 3. Expressions based on the verb *chop* ‘to eat’ in verb-noun combinations.

The above-mentioned expressions are regular structures based on a verb-object pattern.⁵ In most cases, the received effect of combining the verb ‘to eat’ with nouns is a verb. Therefore, *chop* functions as a verbal constituent of phrasal verbs in which the second element determines the meaning (*chop slap* ‘be slapped’, *chop life* ‘to enjoy’).

The use of the verb which means ‘to eat’ as a constituent of phrasal verbs is a characteristic feature of derivational patterns in some West African languages. In Hausa, the verb *ci* ‘eat’ (also *shā* ‘drink’) constitutes the source “of metaphorical mappings into a variety of (sometimes overlapping) semantic domains with basic correlates rooted in real-world physical experiences” (Jaggar & Buba 2009: 230), e.g.: *ci mutunci* (lit. eat decency) ‘humiliate, treat with disrespect’, *ci wākē* (lit. eat black-eyed peas) ‘become pregnant’, *ci kwāf* (lit. eat cup) ‘won cup’. As for *shā* ‘drink’, examples are as follows: *shā wàhàlā* (lit. drink problem) ‘suffer trouble’, *shā rānā* (lit. drink sun) ‘suffer sun’, *shā azūmī* (lit. drink fasting) ‘temporarily break the fast’, *shā kùnū* (lit. drink gruel) ‘frown’ (Newman 2007). Other verbs may also be used in this function, e.g. in Hausa *kashè idō* (lit. kill eye) ‘to wink’, *bugà wayà* (lit. hit a wire) ‘make a telephone call’. Other languages of the area extensively use verb-noun combinations as a derivational pattern in which the verbal meaning ‘do’ is very common, e.g. in Igbo: *gbaa ego* (lit. do money) ‘to pay in’, *gbaa egbe* (lit. do weapon) ‘to shoot up’, *gbaa egwu* (lit. do dance) ‘to dance’ (Pawlak 2010: 196-197).

In Nigerian Pidgin English, the phrases based on a sequence of a verb and a noun in which some other verbs are used also occur. Among them, the verb *get* in combinations with nouns is very common. As most examples with the verb *get* are combined with nouns which denote parts of the body, they will be presented separately later, where the phrases with

⁵ In Hausa, the fixed verb-object collocations are distinguished as idiomatic phrasal verbs (Newman 2000: 260).

parts of the body-related terms will be discussed. Verb-noun combinations with some other verbs are presented below.

Expression	Literal meaning	Translation
<i>carry woman</i>	carry woman	Womaniser
<i>catch trips</i>	catch trips	entertain usually at the expense of another
<i>cry blood</i>	cry blood	Threat
<i>do anyhow</i>	do anyhow	Unruly
<i>do river</i>	do river	well done
<i>drink garri</i>	drink corn flour	in trouble
<i>komot mind</i>	come out mind	to forget
<i>kul tempa</i>	cool temper	to calm down
<i>kwanagida bag</i>	sleep over bag	sleepover bag
<i>shine eye</i>	shine eye	be on [one's] guard/be careful
<i>skata brain</i>	scatter brain	(be) puzzled
<i>watch nait</i>	watch night	security guard

Table 4. Expressions based on other verb-noun combinations.

The presented examples demonstrate that units based on verb-noun combinations are syntactic representation of various grammatical categories: nouns, verbs or and adjectives. The term *adjective* is used here to distinguish the so-called stative verbs which are used to express the “adjectival” meanings. They are equivalents of the English adjectival phrases with the verb ‘to be’, e.g.: *sidonlook* ‘be passive’, *chop cockroach* ‘be pregnant’, *get liver* ‘be brave’, *skata brain* ‘be puzzled’, *do river* ‘well done’. However, the structural pattern of examples from Nigerian Pidgin English is not based on the verb ‘to be’ but on other verbs, and their conceptual basis is different.

The question of grammatical category seems to be irrelevant in many nominal and verbal compounds which are created on the basis of their semantic units rather than their grammatical category. Therefore, the sequence of phrasal components may be different from the examples presented above, i.e. *mama put* (lit. mother put) ‘local restaurants’, *heart cut* (lit. heart cut) ‘be afraid, shocked’.

3.4. Adjective-noun lexical units

The adjectival phrases which include English adjectives are a specific type of expressions which form one lexical unit. The new item has the status of a noun.

Here are examples extracted from the sources:

Expression	Literal meaning	Translation
<i>bad name</i>	bad name	a person with a negative image
<i>big grammar</i>	big grammar	long and difficult English words
<i>brokun plate</i>	broken plate	breakable plates, especially from China

<i>carbon copy</i>	carbon copy ⁶	look alike; replica
<i>monin fud</i>	morning food	Breakfast
<i>native doctor</i>	native doctor	Herbalist
<i>personal person</i>	personal person	best friend
<i>popular side</i>	popular side	cheap seats
<i>pure wata</i>	pure water	bottled water
<i>remote control</i>	remote control	to control a person with voodoo
<i>small chop</i>	small food	Appetizers
<i>strong head</i>	strong head	stubborn; persistent
<i>strong thing</i>	strong thing	something bad
<i>tum-tum belle</i>	round belly	big rounded belly; beer belly
<i>whole man</i>	whole man	Adult
<i>yellow fever</i>	yellow fever	a traffic warden
<i>yellow man</i>	yellow man	Albinos

Table 5. Expressions based on a combination of adjectives with nouns.

Those expressions are assumed to be endocentric compounds⁷ in which the first part of the expressions is the modifier of the head which follows it.

Although in terms of structure expressions from this group follow the rules of Standard English syntax, they are original constructions specific to Nigerian Pidgin English. They are not a copy of expressions which are part of Standard English, therefore they constitute a unique set of adjective phrases. The collocations are copied from native African languages or created *ad hoc* and have a metaphoric provenance. They do not necessarily lead to direct equivalents from African languages, they rather refer to conceptual patterns used to code the particular notion in a particular language. For example, *brokun plate* ‘breakable plates, especially from China’ has the same conceptual basis as the Hausa term *fãdì-kà-mutù* (lit. fall down – die) ‘chinaware, breakable dishes’ (Newman 2007: 57).

The African-based origin of these phrases is also reflected in the violation of the rules of English syntax, when the adjective follows the noun, as in *belle sweet* ‘happiness’, *belle full* ‘satiety (full from eating)’.

The word order in an adjective phrase is specific to a language, and in West African languages the head-modifier sequence varies. In some languages, e.g. in Mande (Creissels 2016: 3f), attributive adjectives are placed in postposition to the noun, in Hausa they may either precede or follow the noun, similarly as in the case of Akan (Appah 2013) in which both types (adjective-noun and noun-adjective) of compounding are possible, i.e. *opanyin bon* (lit. adult bad) ‘irresponsible adult’, but also *asem kese* (lit. matter big) ‘big issue’, *ade tantan* (lit. thing ugly) ‘ugly thing’ (Appah 2013: 77).

⁶ Even though the expression occurs in Standard English it has a different semantic motivation

⁷ **Endocentric compound** – “is the one that has head. The head expresses the core meaning of the compound , and it belongs to the same lexical category as the compound as a whole” (Aronoff, Fudeman 2011:114).

3.5. Lexicalised sentences mirrored from substrate languages

Replacing a word by a sentence or a phrase is a feature of many languages, including Standard English, i.e. *merry-go-round*, *forget-me-not*. In Nigerian Pidgin English this type of phrases is very common and they have different translation equivalents. They have various structural patterns of a clause, very often they are based on an imperative paradigm. However, some phrases presented in the table below are assumed to be exocentric compounds⁸, where there is no head of expression and any part restricts the other parts:

Expression	Literal meaning	Translation
<i>beat am die</i>	beat him die	beat to death
<i>belle-turn me</i>	belly turn me	Diarrhoea
<i>berry don black</i>	berries became black	a way to express that there is enough money
<i>berry never black</i>	berries have not become black	a way to express that you are not wealthy enough yet
<i>bodi du mi</i>	body do me	Premonition
<i>butta my bread</i>	butter my bread	answered prayer
<i>carry am for head</i>	carry it in head	take up too much responsibility for sth; become obsessed with
<i>chuk bodi put</i>	chook body put	to squeeze into a tight corner; getting involved with other people's business
<i>dig am out</i>	dig it out	Fight
<i>draw bodi take</i>	draw body take	Withdraw
<i>ejika ni shop</i>	shoulder is shop (Yoruba <i>ejika ni</i> 'shoulder is')	a tailor walking on the street with his portable sewing machine on his shoulder
<i>eye go come down</i>	eye will come down	come back to reality
<i>go here go there</i>	go here go there	Indecisive
<i>leg no dey comot leg no dey house</i>	leg is not coming out leg	one who goes out too much
<i>leg no dey stay one place</i>	leg is not staying one place	restless; constantly moving about
<i>leg no komot</i>	leg no come out	someone who seems to always be present
<i>mind tell me</i>	mind tell me	Intuition
<i>pe wit tit</i>	pay with teeth	(be) too expensive
<i>put leg for road</i>	put leg on road	start going
<i>take eye see</i>	take eye see	look without touching
<i>wash hand now</i>	wash hand now	come and join me in this meal
<i>your hand reach</i>	your hand reach	to afford

Table 6. List of lexicalised sentences.

⁸ **Exocentric compound** – “compounds whose lexical category or meaning are not determinable from the head” (Aronoff & Fudeman 2011: 114).

Some other examples enrich the list, but they may be qualified as phrases rather than clauses, i.e. *betta follow* ‘good fortune’ (lit. better follow), *money for hand* ‘no credit’ (lit. money in hand), *head no correct* ‘mad; eccentric’, (lit. head [is] no correct), *weak for bodi* ‘in a state of shock’ (lit. weak in body). The two examples: *belle-turn me* and *mind tell me* seem to be a reverted lexicalised form of imperative sentences, where the object has been moved to the first position and it precedes the finite verb and pronominal object.

The lexicalised sentences used as nominal constituents are a characteristic feature of Nigerian Pidgin English. They represent another example of analytical structures and determine the stylistic peculiarities of this language.

This type of desentencisation and the replacing of lexemes by a whole clause is common in West African languages. In Yoruba: *kóiyèsìara* (lit. put mind to body) means ‘to be observant’, whereas *Babalawo* (lit. old man/father has cult) ‘herbalist’ (Taiwo 2014: 50). Examples from Hausa are as follows: *mālām-bùde-littāfi* (lit. Mr open book) ‘butterfly’ (Newman 2007: 149), *fādi – tǎshi* (lit. fall down and get up) ‘struggle, effort’, *dà-na-sanì* (lit. if I had known) ‘regret, repentance’, *kà-cìci-kà-cìci* (lit. make a guess, make a guess) ‘riddle, quiz’ (Piłaszewicz 1990).

4. Parts of the body as a source domain of idiomatic expressions

As constituents of phrasal units, body-related terms occur in different combinations. Regarding the structural patterns, the most common type is represented by combinations with verbs. The second largest group consists of expressions with adjectives. On semantic grounds, the most productive words are *bodi* (‘body’), *belle* (‘belly’), *ai* (‘eye’), *hed* (‘head’), *maut* (‘mouth’). Expressions based on other body parts are also possible, but they are not so common.

4.1. Expressions with *bodi* ‘body’

Following the patterns present in other languages, the term ‘body’ in Nigerian Pidgin English may denote a reflexive concept (Heine 2014: 21), but it had not been grammaticalised into a reflexive pronoun as it happened in other languages of West Africa, i.e. in Yoruba *Nwosu rí ara re* (lit. Nwosu saw his body) ‘Nwosu saw himself’ (Heine 2011: 50).

Other semantic extensions of *bodi* ‘body’ are demonstrated below.

Expression	Literally	Meaning
<i>get bodi</i>	to have a body	Overweight
<i>Gumbodi</i>	to gum a body	to stalk
<i>Joinbodi</i>	to join a body	strike, group
<i>Karibodi</i>	to carry a body	(be) proud
<i>shake bodi</i>	to shake a (one’s) body	to spend money

Table 7. Expressions with the term *bodi*.

The examples indicate that word combinations are seen as units, but the process of their lexicalisation is not complete and the rules of spelling are not consistent. In some cases the

expression consists of two separate words, in some other they are merged. Moreover, the result of this combination may refer to either verbal or nominal constituent, the meaning of *karibodi* '(be) proud' is an equivalent of adjective.

4.2. Expressions with *belle* 'belly/stomach'

Another term which is productively used in the expressions is the word *belle* 'belly/stomach', i.e.:

Expression	Literally	Meaning
<i>belle-turn me</i>	belly-turn me	diarrhoea
<i>get belle</i>	to have a belly	(be) pregnant
<i>get belle to fit laf</i>	to have a belly to be able to laugh	(be) able to laugh
<i>give belle</i>	to give a belly	add volume
<i>hold belle</i>	to hold a belly	prevent hunger

Table 8. Expressions with the term *belle*.

As belly is perceived as the source of emotions in many African languages, Nigerian Pidgin English also shares this feature, but the above-mentioned examples show that the idiomatic expressions involving *belle* may also have other conceptualisations. Most of the expressions are verbal compounds, the compound *belle-turn me* is somehow irregular in this context.

4.3. Expressions with *ai* 'eye'

The table below presents compounds with the word *ai* which in some texts also appear as *eye*, as in the original Standard English spelling. The collected examples are as follows:

Expression	Literally	Meaning
<i>big ai</i>	big eye	ambitious; greedy
<i>chook ai</i>	to chook an eye	to glance
<i>cut ai</i>	to cut an eye	to give an eye signal
<i>difren ai</i>	different eye	opinion
<i>mek ai</i>	to make an eye	to wink
<i>pour ai</i>	to pour an eye	to check out place or something
<i>sharp ai</i>	sharp eye	(be) observant

Table 9. Expressions with the term *ai*.

Collocations which contain the word *ai* are also very common in the language. In most examples *ai* occurs in combination with verbs, therefore *mek ai*, *chuk ai*, *pour ai*, *cut ai*. These are phrasal verbs which are equivalents of English verbs. Only few expressions

are built as phrasal structures with adjectives, i.e.: *difren ai, sharp ai*. They are nominal components of the clause.

4.4. Expressions with *hed* ‘head’

Examples presented in Table 10 are to demonstrate that in Nigerian Pidgin English the word ‘head’ is perceived differently than in many West African languages where ‘head’ is the source of a reflexive pronoun (Heine 2011: 41). For example, in Hausa *Ni da kaina ba na sonta* (lit. I and **my head** do not like her) ‘I don’t like her **myself**’ (Pawlak 2014: 147). Also in Fulani the reflexive pronoun is derived from the lexeme *hoore* ‘head’ (Heine 2011: 57). Semantic derivations of *hed* in Nigerian Pidgin English extend to senses related to the character, state of mind, e.g.:

Expression	Literally	Meaning
<i>bad hed</i>	bad head	unlucky person
<i>get big hed</i>	to have a big head	(be) able to do something
<i>get hed</i>	to have a head	(be) reasonable
<i>trong hed</i>	strong head	Stubborn

Table 10. Expressions with the term *hed*.

Many expressions with *hed* are adjectival phrases, i.e. *big hed, bad hed, trong hed, hot hed*. When used in verbal expressions, i.e. *get hed*, the term ‘head’ is still used to code an adjectival meaning (in the form of stative verb).

4.5. Expressions with *maut* ‘mouth’

The meaning of expressions which contain the word *maut* is differentiated by other words which determine its phrasal collocation. Here are examples extracted from the sources:

Expression	Literally	Meaning
<i>basket maut</i>	basket mouth	(be) blabbermouth
<i>chop maut</i>	to eat a mouth	to kiss
<i>get maut</i>	to have a mouth	talkative
<i>mek maut</i>	to make a mouth	to boast
<i>mek maut like mumu</i>	to make the mouth like a fool	(be) surprised
<i>open mout</i>	to open one’s mouth	to talk
<i>put maut</i>	to put one’s mouth	to comment
<i>sharp maut</i>	sharp mouth	acid tongue
<i>shuk/chuk maut</i>	to shake a mouth	to interfere
<i>smel maut</i>	to smell a mouth	to use offensive words
<i>sweet maut</i>	sweet mouth	sweet tooth; silver tongued

Table 11. Expressions with the term *maut*.

Among the phrases with the word *maut* ‘mouth’, the largest number is represented by combination with verbs, therefore *shuk/chuk maut*, *put maut*, but nominal compounds (*smel maut*, *basket maut*) or adjective phrases (*sharp maut*) are also possible. The idiomatic expression *mek maut like mumu* (lit. make mouth like stupid) ‘to be surprised’ is an instance of African conceptualisation which is based on presenting various states of the human body.

4.6. Expressions with other parts of the body

Many other parts of the body-related terms are used in idiomatic expressions, among them ‘nose’ (NPE: *nose*), ‘face’ (NPE: *face, fes*), ‘hand’ (NPE: *han, hand*), ‘leg’ (NPE: *leg*), ‘teeth’ (NPE: *tit, teeth*), ‘skin’ (NPE: *skin*), ‘throat’ (NPE: *throat*), ‘neck’ (NPE: *neck, nek*), ‘liver’ (NPE: *liver*), ‘heart’ (NPE: *heart, hat*) which occur in the available sources. It is interesting to note that also ‘mind’ and ‘brain’ are used in this function.

Expression	Literally	Meaning
<i>(get) meni leg</i>	to have many legs	complications
<i>back hand</i>	back hand	to slap someone with the back of your hand
<i>blue tooth</i>	blue tooth	to harm (by supernatural means)
<i>cut hat</i>	to cut the heart	to surprise
<i>cut neck</i>	to cut a neck	when you like something so much
<i>drai skin</i>	dry skin	skinny person
<i>drai skin</i>	dry skin	skinny person
<i>draw throat</i>	draw the throat	stimulate appetite
<i>fall hands</i>	fall hands	to act immature/ disappoint
<i>finger foot</i>	finger (of the) foot	toe
<i>get liver</i>	to have a liver	(be) brave
<i>komot mind</i>	come out one’s mind	to forget
<i>long leg</i>	long leg	well connected
<i>long throat</i>	long throat	glutton (Mensah 2011) greedy (naijalango)
<i>mind tell me</i>	mind tell me	intuition
<i>pe wit tit</i>	to pay with teeth	(be) too expensive
<i>poke nose</i>	poke one’s nose	to interpose
<i>raise hand</i>	raise one’s hand	to salute
<i>show 32</i>	to show 32	to smile
<i>shuk leg autsaid</i>	to shook leg outside	to engage in an adulterous relationship by married woman
<i>skata brain</i>	to scatter brain	(be) puzzled
<i>tie face</i>	to tie face	to sulk
<i>troway fes</i>	to throw away face	to ignore

Table 12. Expressions with other parts of the body.

The phrases involving body-part terms represent various structural patterns and refer to many ways of coding the meaning through metaphoric expressions. Along with verbal and nominal phrases, they include lexicalised sentences, such as *mind tell me* ‘intuition’. Particular examples which refer to a specific African conceptualisation draw attention to the use of teeth as a source of other notions, i.e. *pe wit tit* (lit. to pay with teeth) ‘(be) too expensive’, *show 32* (lit. to show 32) ‘to smile’.

5. Concluding remarks: English lexicon – African conceptualization patterns in Nigerian Pidgin English phraseology

Nigerian Pidgin English is very productive in terms of creation of phraseological units which become new lexical items. Even though they are based on the English lexicon, their conceptual basis and structural patterns are copied from indigenous languages.

The examples used to present the structure of idiomatic expressions in NPE confirm that their patterns are found in many African languages of the area. Moreover, the expressions have the same lexical basis. The data allows distinguishing conceptual patterns which they follow in expressing the meaning. From this perspective, the most productive sources of conceptualisation may be distinguished. The collected material justifies the statement that the human body is the richest reservoir of metaphoric expressions which further develop into stable grammatical or lexical units. The examples from West African languages confirm that Nigerian Pidgin English shares the property of using parts of the body-related terms in phraseological units with languages of the area and that their conceptual patterns are similar.

5.1. Verbal phrases and nominal compounds as a source of calques from native languages

Various types of phrases which represent verbal phrases or nominal compounds, as shown in the preceding sections, were also recognized as a strategy for new word formation in a number of West African languages. In many cases, the Nigerian Pidgin English expressions seem to be directly copied from indigenous languages. Whereas the most convincing examples can be taken from Yoruba or Igbo, numerous examples are also found in Hausa, which is not the main source of its structural borrowings,⁹ i.e. NPE *pour ai* ‘to check out a place or something’ is an equivalent of Hausa *zuba ido* (lit. pour eye) ‘to gaze’, while *joinbodi* ‘to gather, strike’ is a structural copy of *hada jiki* (lit. join body) ‘collect in a group’ in Hausa (Pawlak 2014: 145). Similarly *tie face* (lit. to tie face) ‘to sulk’ in NPE resembles Hausa expression for an angry look: *daure fuska* (lit. to tie face) (Almajir 2013: 96). The Hausa compound *yi girman kai* (lit. do big head), which denotes the feature of being ‘full of concepts’ (Almajir 2013: 97) was mirrored in the Nigerian Pidgin English expression *get big hed* ‘(be) able to do something’. Also NPE *sweet mouth* ‘silver tongue’ probably has its source in Hausa *yi zakin baki* (lit. do sweet mouth) ‘present convincing

⁹ See also (Frąckiewicz 2019).

speech' (Almajir 2013: 99). Even if the meanings of some expressions are slightly different¹⁰, mutual inspirations for using the linguistic structures to code the meaning are clear. The phrases *chop cockroach* (lit. eat cockroach) 'be pregnant', *belle sweet* (lit. belly sweet) 'happiness' and *bad-belle* (lit. bad belly) 'malice' are assumed to be direct copies of the Hausa patterns. Even though in Hausa the first example is rendered as *ci wākē* (lit. eat black-eyed peas) 'become pregnant', conceptually it seems to be similar and in both cases it is related to the idea of eating (something). Similarly, expressions with 'belly' might be perceived as conceptually common with Hausa: *farin ciki*² (lit. white belly) 'happiness' and *bākin ciki*² (lit. black belly) 'sadness, envy' (Hollington 2015: 95).

5.1.1. The use of the body-related terms in metaphoric functions

The use of body-part terms to create new expressions is common in many languages of the world. "Body parts are typically animate concepts, being inalienable components of living human or animal organisms. Accordingly, combining these concepts with inanimate possessor nouns constitutes an unconventional process that results in the creation of new concepts" (Heine 2014: 13). Combining the body-related terms with other words brings about the creation of a new compound, phrase, idiomatic expression or grammatical construction (Heine 2014: 13).

In Nigerian Pidgin English, the body-part terms furnish a conceptual basis for phrases which constitute the largest group of metaphoric expressions. In some cases, the same word may be used as a term related directly to a body function or is a metaphoric exponent of a more abstract notion. For example, a lexeme *bodi*, when used metaphorically, denotes a 'person', i.e. *how yo bodi dey* (lit. how is your body) 'how are you', while in other expressions its original meaning 'body' is preserved, i.e.: *get bodi* 'overweight'. The following nouns were identified (Heine 2014; Corum 2015) as a basis for metaphors characteristic of some West African languages.

The metaphoric usage of parts of the body-related terms includes:

- 1) identifying *bodi* with the meaning 'person', as in the greeting formula *How yo bodi de?* (lit. how is your body?) 'how are you?' and its response *bodi dey in cloth* (lit. body is in clothes) 'I'm fine'. Similarly the concept of person/life' is identified with word 'body' in many West African languages. In Hausa the question *Yaya jiki?* (lit. how is body) means 'how are you?' In Ewe "the body and the mind, perceiving, feeling, thinking and knowing, are much closer connected, in fact, inseparable, as compared to the perspectives of the body in the Western culture" (Hollington 2015: 92).
- 2) expressing the notion of SELF with the word meaning 'life', as in *my life don spoil* (lit. my life spoiled) 'I'm destroyed', what can be compared to Hausa: *Raina ya baci* (lit. my life became spoiled) 'I'm sad' (Pawlak 2014: 156). Another expression which confirms this thesis is *Kom giv your life* (lit. come and give your life) 'sacrifice yourself' which could be compared with Hausa: *Ransa yana sansa* (lit. his life wants it) 'he wants it' (Pawlak 2014: 150). This notion has many structural and semantic variants, as demonstrated under 4.5.1.

¹⁰ As an example of semantic shift is the lexeme *raincoat*, which was borrowed from English but in Nigerian Pidgin English it is used to denote the meaning 'condom'.

- 3) perceiving *belly/stomach* as a seat of emotions, i.e.: *belle full* (lit. belly full) ‘satiety’ (full from eating), *belle sweet* (lit. belly sweet) ‘happiness’, *bad-belle* (lit. bad belly) ‘malice’.

According to Hollington (2015), this conceptualisation pattern has a wider African perspective. As opposed to (modern) European languages, where the heart is perceived as a seat of emotions, in many African languages it is rather the belly/stomach which functions as a seat of emotions (Hollington 2015: 94).

That kind of conceptualisation is also found in NPE, where expressions with ‘belly’ are related to emotions, e.g. *Make Nigerians pick candidates wey go sweet their belle* ‘let Nigerians choose candidates who will make them happy’ (lit. make their belly sweet).

The question of how deep these conceptual patterns of expressing the notions are rooted in the West African languages and whether they are shared by other languages of the area, remains the subject of future research.

References

- Almajir, Tijani Shehu. 2013. “The Polysemy of Body Part Terms in Hausa within the Frame of Image Schemas”. *Studies of the Department of African Languages and Cultures* 47, 93-111.
- Appah, Clement Kwamina Insaideo. 2013. “The Case against A-N Compounding in Akan”. *Journal of West African languages* 40 (1), 73-87.
- Aronoff, Mark & Kirsten Fudeman. 2011. *What is morphology?*. United Kingdom: Wiley-Blackwell.
- Corum, Micah. 2015. *Substrate and Adstrate: The Origins of Spatial Semantics in West African Pidgincreoles*. Berlin/Boston: De Gruyter Mouton.
- Creissels, Denis. 2013. “Existential predication in typological perspective”, paper presented at the 46th Annual Meeting of the Societas Linguistica Europaea, Split, 18–21 September 2013, <http://www.deniscreissels.fr/public/Creissels-Exist.Pred.pdf>.
- Elugbe, Ben Ohiomambe & Augusta Phil Omamor. 2007. *Nigerian Pidgin English. Backgrounds and Prospects*. Nigeria: Heinemann Educational Books.
- Faraclas, Nicholas. 1996. *Nigerian Pidgin English Descriptive Grammar*. London: Routledge.
- Faraclas, Nicholas. 2013a. “Nigerian Pidgin structure dataset”. In: Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.). *Atlas of Pidgin and Creole Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology, <http://apics-online.info/contributions/17>.
- Faraclas, Nicholas. 2013b. “Nigerian Pidgin”. In: Susanne Maria Michaelis, Philippe Maurer & Martin Haspelmath (eds.). *The Survey of Pidgin and Creole Languages*, vol 1. Oxford: Oxford University Press, 176-184.
- Frąckiewicz, Olga. 2019. *The Image of African languages structures in Nigerian Pidgin English*, PhD dissertation, University of Warsaw.
- Frąckiewicz, Olga. 2019. “Some common patterns of grammatical structures and linguistic conceptualization in Hausa and Nigerian Pidgin English”. In: Nina Pawlak (ed.). *Linguistic Evidence of Cultural Distance. Hausa in Cross-Cultural Communication*. Warszawa: Elipsa, 194-204.
- Heine, Bernd. 2011. “Areas of Grammaticalization and Geographical Typology”. In: Osamu Hieda, Christa König, Hiroshi Nakagawa (eds.). *Geographical Typology and Linguistic Areas: with special reference to Africa*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 41-66.
- Heine, Bernd. 2014. “The Body in Language: Observations from Grammaticalization”. In: Matthias Brenzinger, Iwona Kraska-Szlenk (eds.). *The Body in Language: Comparative Studies of Linguistic Embodiment*. Leiden: BRILL, 13-32.
- Hollington, Andrea. 2015. *Travelling Conceptualizations. A cognitive and anthropological linguistic study of Jamaican*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

- Jaggar, Philip & Malami Buba. 2009. "Metaphorical extensions of 'eat' =>[OVERCOME] and 'drink' =>[UNDERGO] in Hausa". In: John Newman. *The linguistics of eating and drinking* [Typological Studies of the language, vol 84]. Amsterdam/Philadelphia: John Benjamins Publishing Company, 229-252.
- Matthews, Peter Hugoe. 2014. *The Concise Oxford Dictionary of Linguistics*. Oxford: Oxford University Press.
- Mensah, Eyo Offiong. 2011. "Lexicalization in Nigerian Pidgin". *Concentric: Studies in Linguistics* 37(2), 209-240.
- Mensah, Eyo Offiong. 2012. "Grammaticalization in Nigerian Pidgin English". *Íkala, revista de lenguaje y cultura* 17 (2), 167-179.
- Newman, Paul. 2000. *The Hausa language. An Encyclopedic Reference Grammar*. New Haven: Yale University Press.
- Newman, Paul. 2007. *A Hausa-English Dictionary*. New Haven & London: Yale University Press.
- Pawlak, Nina. 2010. *Języki afrykańskie*. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- Pawlak, Nina. 2014. "The notion of SELF in Hausa". In: Matthias Brezinger & Iwona Kraska-Szlenk (eds.). *The Body in Language: Comparative Studies of Linguistic Embodiment*. Leiden/Boston: BRILL, 140-159.
- Piłaszewicz, Stanisław. 1990. "Phraseologisms in Hausa. New development or an old tradition". *Studies in of the Department of African Languages and Cultures, Hausa Studies III*, 5-21.
- Taiwo, Oye Paul. 2014. "The morpho-syntactic interaction and the derivation of nominal compounds in Yorùbá". *International Journal of Language Studies* 8 (1), 49-71.
- Welmers, William Everett. 1974. *African Language Structures*. Berkeley/Los Angeles/London: University of California Press.

Mary Edward

ORCID 0000-0001-5721-8535

University of Brighton, UK

Lexical iconicity in Adamorobe Sign Language (AdaSL) and Ghanaian Sign Language (GSL)

Abstract

The visual modality of sign languages offers a high potential for iconicity, i.e. resemblance relationships between form and meaning. Of particular interest is the understanding that iconicity is not monolithic but demonstrated through different devices and strategies that may be influenced by cultural factors and specific communicative contexts. This chapter discusses signers and gesturers preference for specific iconic strategies to name handheld tools. Signers of Ghanaian Sign Language (GSL) and Adamorobe Sign Language (AdaSL) were compared with rural gesturers (Adamorobe) and urban gesturers. Working within the cognitive linguistic framework, the chapter discusses signers and gesturers preference for instrument and handling strategies and the consistent use of iconic strategies across signers in each group. Signers and gesturers exhibited systematic preference for iconic representation of tools, choosing an action-based sign depicting how the object is held (handling) or depicting features of the object (instrument). Interesting finding of this chapter is the language contact situation in Adamorobe, and its influence on the rural gesturers' preference for iconic representation although all the gesturers confirmed no prior knowledge of AdaSL.

Keywords: sign language, gesture, patterned iconicity, handheld tools, cognitive linguistics, Ghana, Adamorobe

Introduction¹

Sign languages are the natural languages used by deaf communities all over the world and fulfil all the requirements for full-fledged human language. Sign languages rely on the visual-gestural mode of communication, that is, the signs are perceived by the eyes and delivered by the hands and other parts of the body including the head and face. Even though the signs are produced with the hands, the modality does not determine a universal sign language that is used by all deaf groups and communities all over the world. Different societies, nations and communities have sign languages that are distinct from each other. American Sign Language (ASL) for instance, is different from British Sign Language (BSL) in lexicon and grammar, as well as in the fingerspelling alphabet. Also, within a particular country, there are differences in the sign languages used by deaf communities. In Ghana, for

¹ Portions of this paper are adapted from my ongoing PhD research on Iconicity in Ghanaian Sign Language and Adamorobe Sign Language (Edward forthcoming).

instance, Adamorobe Sign Language (AdaSL), which is an example of a rural sign language, is totally distinct from Ghanaian Sign Language (GSL), which is an urban sign language.

The term “urban sign languages” refer to national sign languages and few other sign languages that are used for deaf education². Urban sign languages are used predominantly by deaf people beside teachers, interpreters (De Vos & Pfau 2015), child of Deaf adults (CODAs), sibling of Deaf adult (SODA) and other hearing members of Deaf families. The term “rural sign languages” refers to sign languages that are used in particular communities with a high incidence of deafness for communication between deaf people and also between deaf and hearing people.

Notwithstanding the difference in modality, sign and spoken languages exhibit similar properties of language structure: i.e. both have phonology, morphology, semantics, syntax etc. (Pfau et al. 2012; Sandler & Lillo-Martin 2006). In recent literature on both spoken and signed languages, *iconicity* (the resemblance relationship between form-meaning) has been noted as a fundamental property of human language (Perniss et al. 2010; Simone 1995). Different frameworks have been used to discuss the notion of iconicity and the most used framework for iconicity in spoken language is the cognitive linguistics framework (Langacker 2008). Other signed linguists have adopted the cognitive linguistics approach to iconicity to discuss iconicity in sign languages (Wilcox 2004; Occhino 2016).

There is a restriction on how much we can represent iconically with sound³ as compared with signs, and this restriction is modality-specific. Gestures and sign languages use the same manual modality for linguistic encoding. The use of co-speech gestures relies on the medium of the hands, face and the body. Similarly, sign language also typically involves the use of articulators other than the hands, notably movements of the mouth. Research on gestures have revealed that gestures share iconic features of visual-spatial modality and exhibit use of similar strategies of iconic representation used by signers (Padden et al. 2015; Padden et al. 2013; Brentari et al. 2015; Ortega & Özyürek 2016). Gesture has the same visual-spatial modality and the expression of information by the hands and other parts of the human body.



1. Patterned iconicity in sign language

Recent linguistic investigations have demonstrated that the ubiquitous influence of iconicity is present in both spoken and sign modality but more productive in sign languages (Occhino 2016; Perniss et al. 2010; Wilcox 2004; Taub 2001; Dingemanse et al. 2015). Iconicity is defined simply as the resemblance relationship between the form and the meaning of a linguistic expression. Valli et al. (2011) state that in linguistic iconicity, the form of the symbol is an icon or picture of some aspect of the thing or activity being symbolized and that the way of representing this iconicity is sign language specific. For example, the sign HOUSE⁴ in GSL is a picture of the roof of a building and the sign WOMAN in AdaSL

² In most parts of the world sign languages have only quite recently been used in education.

³ Occhino states that the “mismatch in our ability to iconically represent our interactions in space and time, through sound, is compounded by the fact that the oral channel is much more restricted in terms of articulatory degree of freedom and visual saliency” (Occhino 2016:193).

⁴ The glossing convention for sign language use capital letters for signs.

is iconic of a woman's breast. The resemblance relationship here is the mapping of the mental representation of the articulatory forms represented as  for HOUSE and  (touching a meaningful location, i.e. the chest) for WOMAN. The linguistic representation of the articulatory form according to P. Wilcox (2000) has a resemblance relationship with the mental depiction of the concept in an individual's experience. Different types of iconicity exist in different sign languages and these include handling (holding or manipulating), object (hand as object), instrument (hand as object in use), etc. Signers employ the diverse possibilities afforded by the hands and the body to represent iconic structures in sign languages (Meir et al. 2013).

The recurrent use of an iconic strategy across signs in a particular semantic category has been called Patterned Iconicity (Padden et al. 2013; 2015). Different sign languages exhibit different patterns of iconicity within a category (Kimmelman et al. 2018). For example, Padden and her colleagues found differential patterns with respect to the use of *handling* or *instrument* strategies for the semantic category of handheld tools by signers and gesturers⁵ (American Sign Language (ASL), New Zealand Sign Language (NZSL), Al-Sayyid Bedouin Sign Language (ABSL), American gesturers and Israeli Bedouin gesturers) (Padden et al. 2013). They identified a preference for handling strategy by the gesturers: i.e. the hand holding or grasping an object. For signers, ASL and ABSL signers favoured instrument strategy: i.e. the hand depicts features of the object and performs canonical actions of the object. NZSL signers favoured the handling strategy to name handheld tools (Padden et al. 2013).

Kimmelman et al. (2018) identified iconicity patterns for different semantic categories in their Iconicity Patterns in Sign Languages (IPSL) database. They identified different iconic strategies for the semantic category named Instrument. They identified British Sign Language (BSL), French Sign Language (LSF), German Sign Language (DGS) among others as *handling* in relation to the semantic category of instrument. American Sign Language (ASL), Estonian Sign Language and Brazilian Sign Language preferred the *object* strategy according to Kimmelman et al.'s analysis. On the other hand, Russian Sign language and Italian Sign Language used both *object and handling* strategies to the same degree.

2. Cognitive Linguistic Approach to Sign Languages

In the Cognitive Linguistics (CL) framework language is assumed to reflect fundamental properties and design features of the human mind (Evans & Green 2006). In CL, linguistic symbols (linguistic units) have both phonological and semantic poles and the relationship between the poles could be that of resemblance or arbitrary (Langacker 2008; Wilcox 2004). For example, in GSL and AdaSL, there is a resemblance relation between the form and meaning of TABLE. The resemblance is the flatness that is represented by both sign languages. This iconic relationship between the form and the meaning is represented by the phonological pole (handshape- palm facing down or palms tracing the flat surface) and

⁵ Gesturers in this chapter refers to hearing people who used hand and body gestures to name the handheld tools.

the semantic pole (meaning- a piece of furniture with a flat top). However, when there is no resemblance relation between the phonological and semantic poles, there is no iconicity.

A model of iconicity formulated for sign language and relevant for the chapter is *Cognitive Iconicity* developed by Sherman Wilcox (Wilcox 2002; 2004). Wilcox defines cognitive iconicity as “a distance relation between the phonological and semantic poles of symbolic structures” (Wilcox 2004:122). In the cognitive iconicity perspective, arbitrariness is reduced because “the phonological and semantic poles of signs reside in the same region of conceptual space” (Wilcox 2004:122). The phonological pole refers to the linguistic form and the semantic pole refers to the actual meaning given to the word (see figure 1). The form entails the phonological realisation of the sign⁶; Handshape (HS), Movement (M), Location (L) and Orientation (O) and these phonological parameters are relevant for the meaning of the sign. The Handshape is the acceptable hand form in signing; Movement refers to how the dominant or the non-dominant hand or both are moved to create a sign; Location is the place the dominant hand or the non-dominant hand is placed in the formation of a sign and Orientation is direction of the palm of the dominant hand.

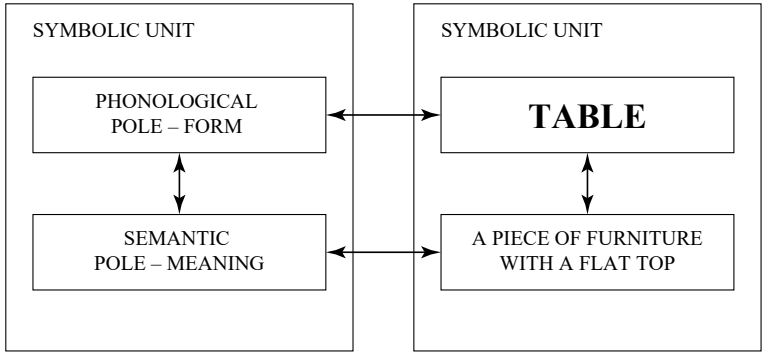


Figure 1. Phonological and semantic poles

The phonological pole, which consists of the articulatory parameters, represents the *construals* of form and the semantic pole (which can have arbitrary or iconic relationship with the phonological pole) represents the *construals* of meaning. Langacker (2008:43) defines *construal* as “our manifest ability to conceive and portray the same situation in alternate ways”, and as such TABLE can be construed as *flat object* or *four-legged entity* depending on what is profiled by signers.

⁶ The phonology is primarily made up of the articulatory parameters and the non-manual markers. The articulatory parameters refer to the aspects that come to play in the formation of the sign. These include the Handshape, Movement, Location and Orientation. The non-manual markers in sign languages rely on facial expression and the position of the body and these act as prosodic cues to the sign.

The phonological parameters are morphological in other parts of the lexicon, such as classifier constructions, that is a signed word that represent a group of objects and is sometimes incorporated with movement (Occhino 2016).

3. Language Information

Ghanaian Sign Language (GSL) is the sign language of the urban deaf community in Ghana. There are 110,625 deaf people in Ghana⁷ that is 0.4% of the population and out of these number, some people use GSL as either a first or second language. GSL developed from the sign language introduced by Andrew Foster⁸ in 1957 and it is representative of Ghanaian society and Ghanaian culture. Most urban deaf people are educated in GSL. It is used by the state television to interpret news items and special national events to deaf people and other users of GSL. The first dictionary for GSL is believed to have been produced in the early 2000s and a new dictionary edited by Deutsch & McGuire was published in 2015 (McGuire & Deutsch 2015)⁹.

Adamorobe Sign Language (AdaSL) is an indigenous village sign language used in Adamorobe in the south eastern part of Ghana. AdaSL is believed to have existed as far back as 1733 as a language used by both hearing and deaf people in Adamorobe (Okyere & Addo 1994). The community is noted for its unusually high incidence of hereditary deafness of an estimated 40 deaf people in the village (1.3% of current population of about 3000)¹⁰. Miles (2004) reported that deaf Adamorobeans were the first substantial historical group of African people known to have used a formal sign language and the record dated as far back as the 18th century. The sign language in Adamorobe is older than GSL and has a long tradition of usage by both deaf and hearing people. Earlier research done in Adamorobe discovered that almost everybody in the village could communicate in the sign language (Frishberg 1987). However, current visits to the community reveal a decline in the numbers of hearing and deaf signers. Linguistically, AdaSL is distinct from GSL and other home signs used in Ghana.

4. Methods

4.1. Participants

Four groups of signers and gesturers were recruited for this study. For the sign languages, 10 signers of GSL were recruited from Medie and Nsawam (Greater Accra and Eastern Regions) and 10 signers of AdaSL were recruited from Adamorobe in the Eastern region of Ghana. The groups of gesturers or non-signers were recruited from Adamorobe (n=4) and Sekondi (n=6) representing rural and urban locations respectively.

All deaf signers were deaf adults between the ages of 24 and 65 years and they were recruited based on willingness to join the study¹¹. Gesturers were all confirmed by

⁷ Accessed from Ghana National Association for Deaf (GNAD) website 31/08/2018. GSL is mostly used by educated deaf people.

⁸ Andrew Forster was a deaf African American who contributed immensely to Deaf education in Sub-Saharan Africa (Kiyaga & Moores 2003).

⁹ A free online mobile application for GSL has recently been launched by Leiden University's Lab for Sign languages and Deaf studies (2020).

¹⁰ Nyst cites 2% of population of 2400 in a research done in 2001 (Nyst 2007).

¹¹ All deaf participants were given detailed explanation of the research work and they all consented by signing a consent form. They were also paid for taking part in the research. All non-signers consented to

friends and family as having no sign language knowledge and were between the ages of 19-57 years. Gesturers from Adamorobe were specifically scrutinized to ensure that they had no knowledge in AdaSL¹².

4.2. Data collection process

All data collection involved picture elicitation task. Pictures of 20 handheld tools were displayed on a laptop screen and signers and non-signers were tasked to name the object in their sign language or with a gesture. The stimulus materials were displayed on white background and stayed on the screen until the signer or gesturer had finished producing the sign or gesture for it.



Figure 2. Examples of handheld tools used for elicitation

Signers and gesturers were first shown the materials and were given the opportunity to look through the pictures before the elicitation task. All AdaSL signers signed in the presence of the researcher and the camera at different times. 40% of GSL signers signed before the researcher and the camera. 60% signed in the presence of the researcher and a research assistant (non-signing). All urban non-signers gestured in the presence of the researcher and they all signed on different days and different times. Rural non-signers gestured in the presence of the researcher and two hearing signers of AdaSL. The presence of the AdaSL signers was to ensure that no non-signer had knowledge of AdaSL.

4.3 Coding

The coded handheld tools are *bottle, broom, bucket, comb, cup, fork, hammer, iron, key, knife, lipstick, long broom, mobile phone, paintbrush, pen, saw, scissors, spoon, toothbrush, and umbrella*. The 20 handheld tools resulted in 251 tokens from the 10 GSL signers, 241 tokens from the 10 AdaSL signers, 187 tokens from the 6 urban signers and 103 tokens

take part in the research work, their participation was voluntary and they were not paid for taking part in the research.

¹² Although earlier research in Adamorobe concluded that almost everyone (deaf/hearing) could sign (Frishberg 1987), recent research has revealed the gradual decline in the number of hearing signers (Edward 2015). There are more non-signing hearing people in Adamorobe than hearing signers.

from the rural gesturers. The responses by deaf signers consisted of individual signs and a few multipart signs. Rural gesturers mostly responded with single tokens and very few two-part signs. Items that elicited more than one sign mostly used different strategies for each sign. For example, BOTTLE (figure 3) elicited handling and entity strategies for two-parts signs (AdaSL), and elicited strategies like tracing 3D (GSL) and handling (gesturer).



Figure 3. BOTTLE by signers and gesturers

Coding was done with the language annotation software ELAN (version 5.4, 2018) (Wittenburg et al. 2006) developed at the Max Planck Institute for Psycholinguistics. The videos were coded for use of iconic strategies listed below:

(a) *Handling* (Padden et al. 2013; 2015). In this strategy, the hand(s) represents human hand holding or grasping object. The hand holding or grasping the object can also perform canonical actions related to the object. With the *hand as hand* schema, “the entire hand is profiled, and is construed as a hand” (Occhino 2016:144).

(b) *Instrument* (Padden et al. 2013; 2015). In this strategy, the hand(s) depicts features of object and performs canonical actions related to the object. The Handshape (HS) has some or all the features of the depicted object. Instrument schematised the handshape as *object-shape schema* (Occhino 2016) where the hand is an entity (or object) performing a canonical action related to the object (instrument) as in mobile phone. When the hand only shows features of the object and does not perform any action, we get the entity strategy.

(c) *Entity*- This strategy is referred as the *object* strategy by other researchers (Padden et al. 2013; Kimmelman et al. 2018; Hou 2018). The hand only shows features of object and does not perform any action with this strategy. Another type of entity strategy that has been identified in this work is *Entity at body Location*. The entity handshape located at a meaningful location shows features of the object but in addition has a meaningful location. The hand at the location does not perform any action.

(d) *Measuring* (Mandel 1977; Ebling et al. 2015; Nyst 2016). Hand(s) indicates size or the height of entity, or the size of the object is shown by delimiting relevant part of the finger, hand or arm. Nyst refers to this strategy as measure stick.

(e) *Tracing (2D)* – For this strategy, the hand(s) are drawing tool and the movement creates virtual shape in space or on body. This result is 2D shapes. It is referred to as sketching in other literature (Mandel 1977; Ebling et al. 2015).

(f) *Tracing (3D)*- The hands and movement represent shape of entity by tracing outline or surface of entity. This results in 3D shapes (Ebling et al. 2015; Padden et al. 2013; Kimmelman et al. 2018).







(g) *Indexing*- Signers' hand points to objects at present or point (or hold on) to part of signer's body that is related to the object.

(h) *Body strategy*- Parts of the body perform canonical actions related to it. The body strategy is distinct from the handling strategy. Body strategy includes shrugging shoulder, a leg kicking, body reclining, arms moving up and down among others.

(h) *Presentable Action*- This refers to embodied conceptual gestures in a source culture that are used to code experiences (both abstract and tangible) or to name items that are related to the experiences generated by an object. For example, the gesture of smelling roses- FLOWER; the gesture of putting head down on pillow- SLEEP.

A few of the signs and gestures were also coded as *not clear* in their iconicity strategies used. These signs on the surface were perceived as having some form of iconic form-meaning mapping. However, the mapping between the form of the sign and the referent was not clearly marked as one of the above listed strategies. For example, initialized¹³ signs (used by GSL signers) that had meaningful locations and (or) meaningful movement but did not have a specific iconic strategy were marked as *not clear*. The responses marked as *not clear* did not clearly fit with the iconic strategies listed above. There were few non-iconic signs from the signers (no form-meaning resemblance relationship).

Below are a few examples of the strategies as used by signers and gesturers to name handheld tools.

Iconicity Type	Examples		
Handling	 <p data-bbox="337 1134 454 1157">KEY (AdaSL)</p>	 <p data-bbox="589 1134 801 1157">LONG BROOM (AdaSL)</p>	 <p data-bbox="822 1134 939 1157">COMB (GSL)</p>
Instrument	 <p data-bbox="337 1457 549 1479">TOOTHBRUSH (AdaSL)</p>	 <p data-bbox="602 1457 807 1479">MOBILE PHONE (GSL)</p>	 <p data-bbox="908 1457 1040 1479">SAW¹⁴ (AdaSL)</p>

¹³ Initialization in sign language is the process of using the fingerspelled letter of the alphabet that represents the first letter of the English word as the handshape for a sign. The initialized handshapes did not have iconic mappings as described above.

¹⁴ This sign has figure-ground relationship which will be discussed in section 7.


<p>Entity</p>			
	MOBILE PHONE (GSL)	FORK (AdaSL)	MIRROR (GSL)
<p>Measure¹⁵</p>			
	BOTTLE	(AdaSL) CUP (AdaSL)	
<p>Tracing 2D¹⁶</p>			
	BOTTLE (rural gesturer)	BUCKET (urban gesturer)	CUP (urban gesturer)
<p>Tracing 3D</p>			
	BOTTLE (GSL)	CUP (GSL)	BUCKET (urban gesturer)

Table 1. Iconic strategies with examples

¹⁵ In these examples, the upright hand is the entity handshape and the index finger is the measure stick that delimit the relevant part of the hand to depict BOTTLE and CUP.

¹⁶ In tracing 2D, the hand is a drawing tool that traces the shape of the object. It is mostly depicted with the index finger. On the other hand, tracing 3D is mostly depicted with the open palm.



Not clear & Not iconic	Examples		
Not clear	<div style="display: flex; align-items: center;">  <div style="margin-left: 20px;"> <p data-bbox="698 366 975 455">In this sign, the movement is meaningful, but the HS is initialised from KEY.</p> </div> </div> <p data-bbox="294 557 429 578">KEY with K HS</p>		
Not iconic	<div style="display: flex; justify-content: space-around;">  </div> <p data-bbox="294 924 524 945">SAW (fingerspelling) (GSL)</p>		

Table 2. Examples of signs marked not clear and not iconic

The responses from the signers (not the gesturers) were also coded for consistency and agreement in the use of iconic strategies across signers in each group.

5. Results

The main iconic strategies that emerged were the handling and instrument strategies for naming handheld tools by both signers and by the gesturers. Overall, GSL and AdaSL signers used more instrument strategy as compared to the gesturers. Although handling strategy was preferred by gesturers (rural/urban), rural gesturers showed higher preference for instrument as compared to urban gesturers. Other strategies were used minimally by both signers and gesturers.

5.1. Iconicity patterns among signers

Among GSL signers, an average of 53% used an instrument strategy and 30% used a handling strategy. Other strategies used were the tracing (3D) strategy (8%) and entity strategy (4%), and 9% of the GSL signs for handheld tools were coded as *not iconic* (see Table 2 for example).

AdaSL signers, just like their urban counterparts, demonstrated a higher preference for an instrument strategy in naming handheld tools and objects. 58% of the AdaSL responses reflected an instrument strategy, 26% of responses were a handling strategy and 8% of the responses used entity strategy. There was no response in the AdaSL data that used tracing (2D & 3D) and the body strategy was not used by GSL signers. Only 3 tokens from AdaSL were coded as non-iconic; one was a borrowed sign from GSL (WATER).

Comparing the iconic strategies used by GSL and AdaSL signers to name handheld tools, the overall pattern is the use of instrument and handling strategies by signers (figure 4) as demonstrated in previous studies (Padden et al. 2013; Padden et al. 2015; Kimmelman et al. 2018).

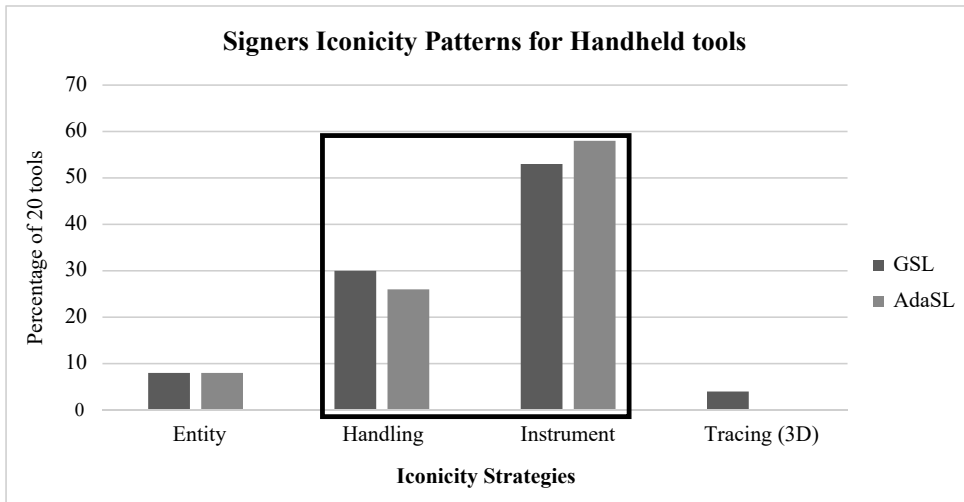


Figure 4. GSL and AdaSL iconicity patterns for handheld tools¹⁷

In both sign languages, there were items that elicited more *instrument forms* and these included broom (100% GSL, 100% AdaSL), comb (80% GSL, 100% AdaSL), fork (100% GSL, 100% AdaSL), mobile phone (100% GSL, 100% AdaSL), paintbrush (100% GSL, 100% AdaSL), saw (90% GSL, 100% AdaSL), scissors (100% GSL, 100% AdaSL), spoon (100% GSL, 100% AdaSL) and toothbrush (80% GSL, 100% AdaSL). Some tools elicited more *instrument forms* in one sign language as compared to the other. Examples include knife (10% GSL, 100% AdaSL) and pen (70% GSL, 0 AdaSL). There were few items that elicited more *handling forms* in both sign languages. These include the following: bucket (100% GSL, 90% AdaSL), hammer (80% GSL, 100% AdaSL), iron (90% GSL, 80% AdaSL), key (80% GSL, 100% AdaSL), long broom (70% GSL, 60% AdaSL) and

¹⁷ The differences in the preference for either instrument or handling among GSL and AdaSL signers did not show significance. GSL and AdaSL signers used instrument and handling forms almost to the same degree. For example, the preferences for instrument forms; AdaSL (58% mean) signers over GSL signers (53% mean) in relation to 20 handheld tools did not show statistical significance (Two-sided Wilcoxon Rank W = 32, p-value = 0.1807). The few numbers of participants made statistical inferences difficult to run for the gesturers.

umbrella (70% GSL, 70% AdaSL). Other tools elicited more *handling forms* in one sign language as compared to the other: lipstick (80% GSL, 20% AdaSL), pen (10% GSL, 100% AdaSL). Few items had high responses for using other iconic strategies in one or both sign languages and these include bottle (tracing (3D)-100% GSL); bottle (entity, 70% AdaSL) and umbrella (entity 50% GSL, 40% AdaSL).

5.2 Iconicity patterns among gesturers

Gesturers from urban and rural locations also made predominant use of handling and instrument strategies for the category of handheld tools (figure 5). Although most of the responses from the gesturers used handling strategy as seen on other studies (Brentari et al. 2015; Ortega & Özyürek 2016; Padden et al. 2013), gesturers from Adamorobe (rural) showed a higher preference for instrument strategy in comparison with their urban counterparts and similar to what was demonstrated by the AdaSL signers. 62% of the average responses of the urban gesturers used handling strategy and 51% of the average responses of the rural gesturers were in the handling strategy. However, rural gesturers used instrument strategy to a higher degree (42%) as compared with urban gesturers (23%).

There were varied responses by gesturers. For instance, although all rural gesturers responded with an instrument strategy for *knife*, only three out of the six urban signers responded with an instrument strategy. All rural gesturers responded with instrument strategy for *broom* whereas urban gesturers responded with a handling strategy. Further, *toothbrush* had varied responses from both gesture groups; rural (3 handling, 1 instrument) and urban (4 handling, 2 instrument).

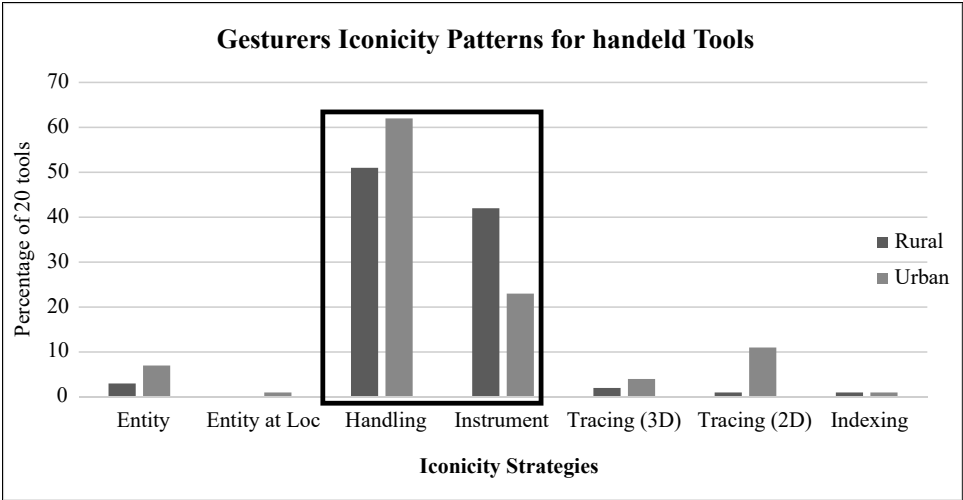


Figure 5. Rural and Urban gesturers iconicity patterns for handheld tools

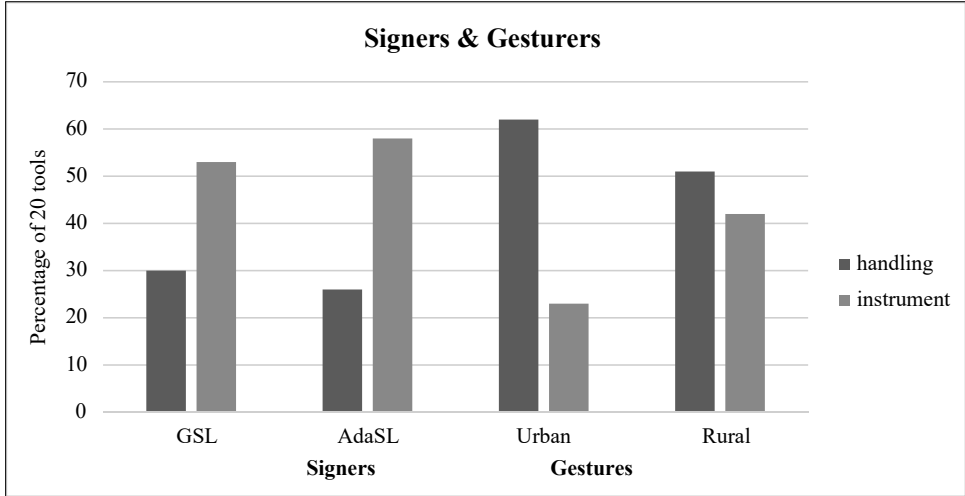


Figure 6. Signers and non-signers compared¹⁸

5.3 Consistency and full agreement among signers

Consistent use of strategy is defined as >70% of signers use the same strategy (Pad-den et al. 2013) and full agreement is defined as 100% of signers use the same strategy (see figure 7). AdaSL signers had 90% consistent use of strategies and GSL had 85% of consistent use of strategies. AdaSL signers had 65% full agreement of strategy and GSL signers had 45%. Consistency and agreement in use of same strategy for individual items is higher in AdaSL than in GSL

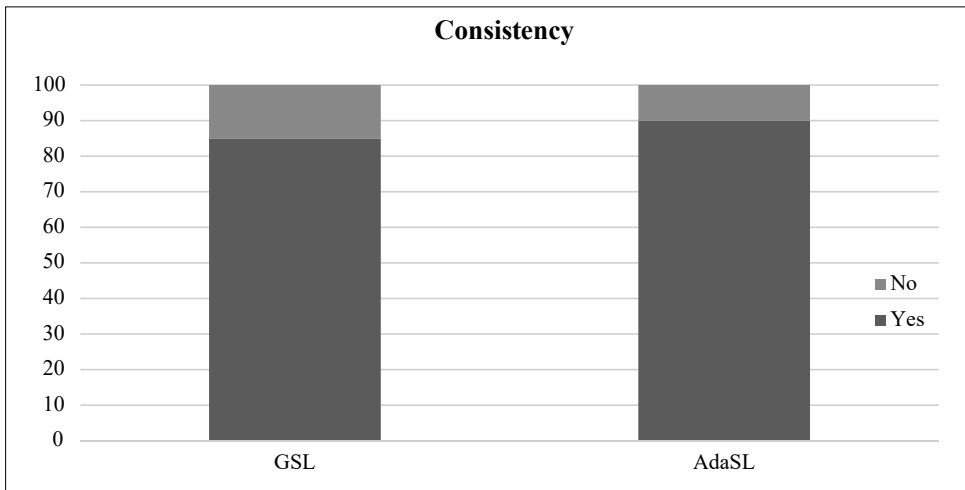


Figure 7. Consistent use of strategy by GSL and AdaSL signers

¹⁸ In this chart, black depicts handling strategy and grey depicts instrument strategy (across signers and gesturers)

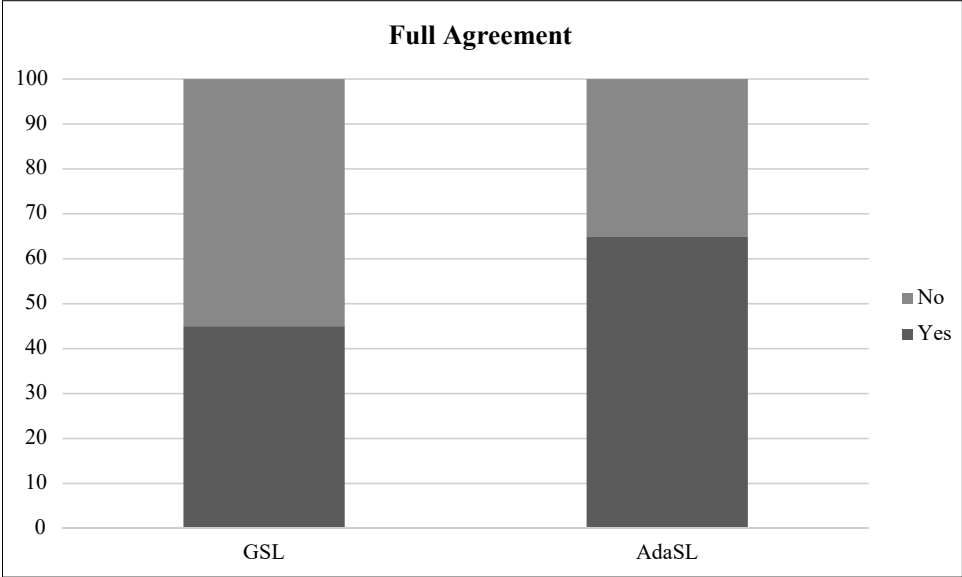


Figure 8. Full agreement in use of strategy



Figure 9. BROOM – consistent use of strategy by AdaSL signers; 100% instrument

6. Discussion

Signers and gesturers make predominant use of handling and instrument strategies for the category of handheld tools. Both GSL and AdaSL signers prefer the instrument strategy. Urban gesturers prefer the handling strategy. The preference for handling forms by gesturers could indicate that gesturers prefer to name instrument forms according to how they are handled whereas signers present extra information that depicts features of

the object. Therefore, in agreement with Padden et al. (2013) preference for instrument forms seems to emerge quickly in a new sign language like ABSL. GSL, which emerged in the 1950s, demonstrated a high preference for instrument forms. Padden et al. (2013) also links the preference for the use of handling strategy to be associated with the use of mouthing as a way of marking nouns as exemplified with the prevalent mouthing in New Zealand Sign Language (NZSL). Mouthing in sign language refers to the oral production of the syllables in the word i.e. signers say the word as used in the equivalent spoken language at the same time of sign production. German Sign Language (DGS) and British Sign Language (BSL) (Kimmelman et al. 2018) prefer the handling strategy as compared to instrument and this could also be linked to the prevalence of mouthing in both sign languages. Both GSL and AdaSL signers used little mouthing.

Whereas instrument strategy permits diverse construed variants, handling strategy has little allowance for profiling the hand-as-hand with construed variants. For example, construed variants of the *mobile phone* profiled different handshapes but used the same iconic strategy. The similarity between the linguistic form of *mobile phone* and the meaning is the representation by the construals of handshape¹⁹ represented by signers. In representing *mobile phone*, the iconic relationship between the form and the meaning is represented by the phonological pole (i.e. representing *mobile phone* with different handshapes depending on the how signers profiled the depicted image). On the other hand, when we cannot match the relation between the phonological and semantic poles, there is no iconicity. From the example of *mobile phone* given in figure 10, we can see that the meanings of the lexical signs are construed based on signers’ perception of the real world and profile of the object.






Iconic strategy	Construed variants using the same iconic strategy			
Instrument 	 GSL(Y-HS) MOBILE PHONE	 GSL(G-HS)	 AdaSL (A-HS)	 AdaSL (B-HS)

Figure 10. Variants of mobile phone

Although the instrument strategy uses an entity handshape²⁰, the entity strategy resulted in 8% of the responses from both GSL and AdaSL signers. Even gesturers from Adamorobe showed higher preference for instrument in comparison to urban gesturers. For example, all rural gesturers responded with instrument strategy for *broom* (just like rural signers, see Figure 9) whereas urban gesturers responded with a handling strategy. The high proportion

¹⁹ The handshapes are named according to their closeness with the alphabets (GSL alphabets). See the examples in fig.10. See Appendix for the GSL alphabet.

²⁰ In the instrument strategy, the hand(s) depicts features of the object (entity) but in addition, the entity handshape performs canonical actions related to the object. On the other hand, the entity strategy only shows the features of the object and does not perform any action.

of use of instrument strategy by gesturers in rural Adamorobe community may be the result of language contact situation with AdaSL²¹. However, because the number of gesturers was lower than the signers (urban=6, rural=4), we need to treat this result with caution because the differences could be a matter of chance.

In cognitive linguistics, the relationship between the linguistic form and the meaning is inseparable and language reflects the properties of the mind. Therefore, the linguistic form has a psychological reality drawn from the experiences of a language user. In the analysis of lexical items from signers (and gesturers), the phonological forms of the signs are emergent units of constant use and that reflects the properties of the mind. The handshape of the sign bears semantic features of the object based on association. Different handshapes profile different semantic features of the object: profiling the whole object or profiling part for the whole (metonymy). The same object (such as *broom* and *mobile phone*) can be construed in alternate ways and each alternate way will profile particular semantic information.

Language users perceive linguistic content in relation to their understanding. The relationship between the form and the meaning is representative of how we represent conceptual structures with linguistic forms and how the forms resemble the structures they represent. Based on the above data, signers of GSL and AdaSL mostly construed handheld tools with instrument or handling strategies. Hwang et al. (2017:10) refers to these two strategies jointly referred as the *manipulation strategy* because they involve “the body representing the body of a human agent and an arm representing the arm of a human agent as it acts upon the referent”. Instrument strategy has entity handshapes with action-associated movement and the movement of the hand/arm represents the action of the agent on the referent.

Also identified in the data is that construal of the same strategy could also have different phonological profiling depending on signers’ perception of objects. For example, the same strategy can have change in HS, location, movement or orientation depending on how signers conceptualize the articulators. From the data, all the 10 GSL signers profiled *mobile phone* as an instrument, 8 of the signers used the Y-HS focusing on the shape of telephone and 2 others (marked with a black box) used the G-HS focusing on the *mobile phone* with a pole (see figure 11). All AdaSL signers also profiled *mobile phone* as an instrument, 9 of the signers used the A-HS and one signer profiled it with both a B and G handshapes. That is, both GSL and AdaSL signers construed mobile phone as an object and focused on different shapes of the mobile phone. Whereas the Y-HS depicted the older version of telephone and the G-HS depicted a phone that has a pole, the A-HS & B-HS depicted a little handheld or portable device. In other words, the different handshapes refer to different types of phones as profiled by the signers.

In cognitive iconicity (Wilcox 2004) the phonological and semantic poles reside in the same region of conceptual space. In other words, the linguistic form of the lexical signs and the meaning are embedded in the same mental concepts. The iconic strategies used by signers and gesturers encompassed the phonological and semantic poles through profiling of the object. The phonological parameters (Handshape, Movement, Location and Orientation) and the semantics (meaning) contribute to iconic structures. The phonological

²¹ The rural gesturers live together in the same community with AdaSL signers and the imperatives of day to day conversations might trigger the visibility of some AdaSL signs to these nonsigners.

and semantic poles of *mobile phone* in GSL and AdaSL have a resemblance relationship²²; that is, all handshapes profile *mobile phone* depending on how signers construe the object with the iconic strategies. The iconic handshapes share form-meaning mappings with the construals of the linguistic form and the construals of real-world objects. AdaSL signers demonstrated greater consistency of strategy because >70% of the tools were profiled with the same iconic strategies by the signers and there was 65% full agreement of strategies used (see Figures 7 & 8). AdaSL as village sign language with long history shows higher consistency and agreement in use of strategy compared to younger GSL. GSL is relatively new sign language (that emerged in 1957).



Figure 11. MOBILE PHONE, GSL

The phonology of *mobile phone* and most of the lexical signs discussed above have meaning and this meaning is associated with the semantic pole (see figure 12). In consideration, the phonological pole is submerged in the semantic pole of a symbolic unit (Langacker 2008; Wilcox 2004; Occhino 2016). In this sense, the different representations of the phonological pole by signers (and gesturers) are the result of different usage-events that are triggered by the object.

Signers handshape “exists on a continuum of form-meaning mappings ranging from near one-to-one mappings to many-to-one mappings” (Occhino 2016:4). One-to-one mappings deal mostly with the phonology and many-to-one mappings present morphological information (Occhino 2016). Unlike the one-to-one mapping that elicited single forms, the hands profiled different information (different forms) based on many-to-one mappings. There were some handshapes that were higher in the number of times used than others. For instance, the B-HS resulted in 50 responses out of the 252 tokens of GSL signers and the A-HS resulted in 36 tokens. In AdaSL, the B-HS resulted in 77 responses of the 242 tokens of AdaSL signers and the A-HS resulted in 42 responses. The higher usage of B-HS and A-HS by both GSL and AdaSL signers predict a higher number of associations with the handshapes, that is many of the lexical items were profiled with B and A handshapes. On the other hand, F-HS elicited 10 (out of 252 tokens) responses from GSL and 7 (out of 242 tokens) from AdaSL signers; lower associations and frequency show that there were fewer items that profiled the F-HS.

²² There is resemblance relationship when the phonological and semantic poles are close together.

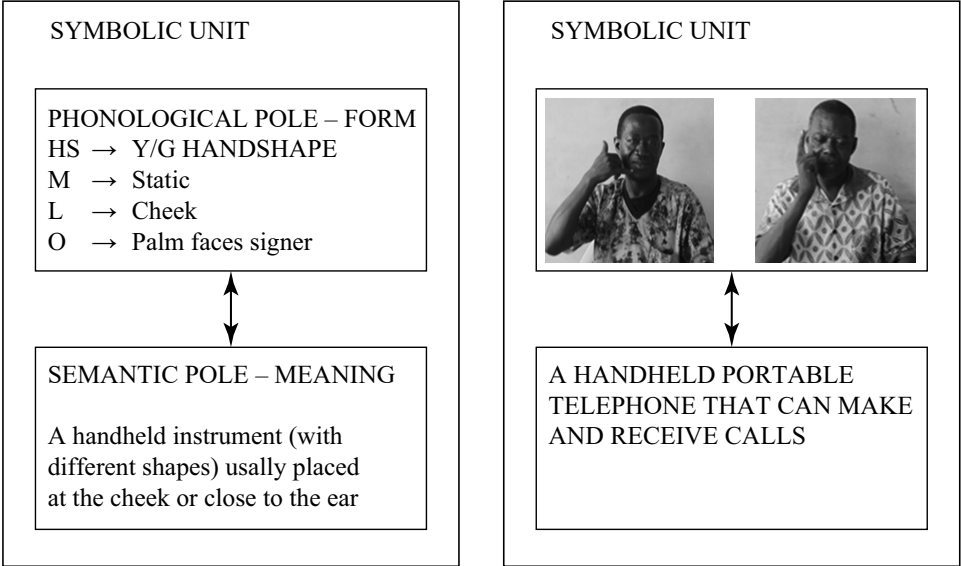


Figure 12. Phonological and semantic poles

Finally, few of the responses had figure-ground relationships in individual signs. For these responses, each hand – the dominant hand and the non-dominant hand – conveyed different aspects of the object. For most of the signers, the dominant hand profiled the figure whereas the non-dominant hand profiled the ground (or the entity being acted on). the dominant hand for all signers shown in figure 13 is the right hand.

<i>Examples from GSL</i>		<i>Examples from AdaSL</i>	
FORK	KEY	KNIFE	HAMMER

Figure 13. Examples of signs with multiple schemas

Lexical items that elicited multiple forms had thematic roles like agent, instrument and patient (see table 3). The dominant hand (for most of the signers the right hand) acted upon the non-dominant hand in agent-patient relationship. For example, in figure 13 the dominant hands were fork (instrument), key (handling), knife (instrument) and hammer (handling). The non-dominant hands elicited entity strategy.

Articulatory unit	KNIFE	HAMMER	FORK	KEY	Thematic role
Dominant hand	B	F-closed	H	A	Agent
Non-dominant hand	B	A	B	B	Patient
Iconic strategy (dominant)	Instrument	Handling	Instrument	Handling	Agent/instrument
Iconic strategy (non-dominant)	Entity	Entity	Entity	Entity	Patient

Table 3. Multiple schemas in individual signs.

7. Conclusion

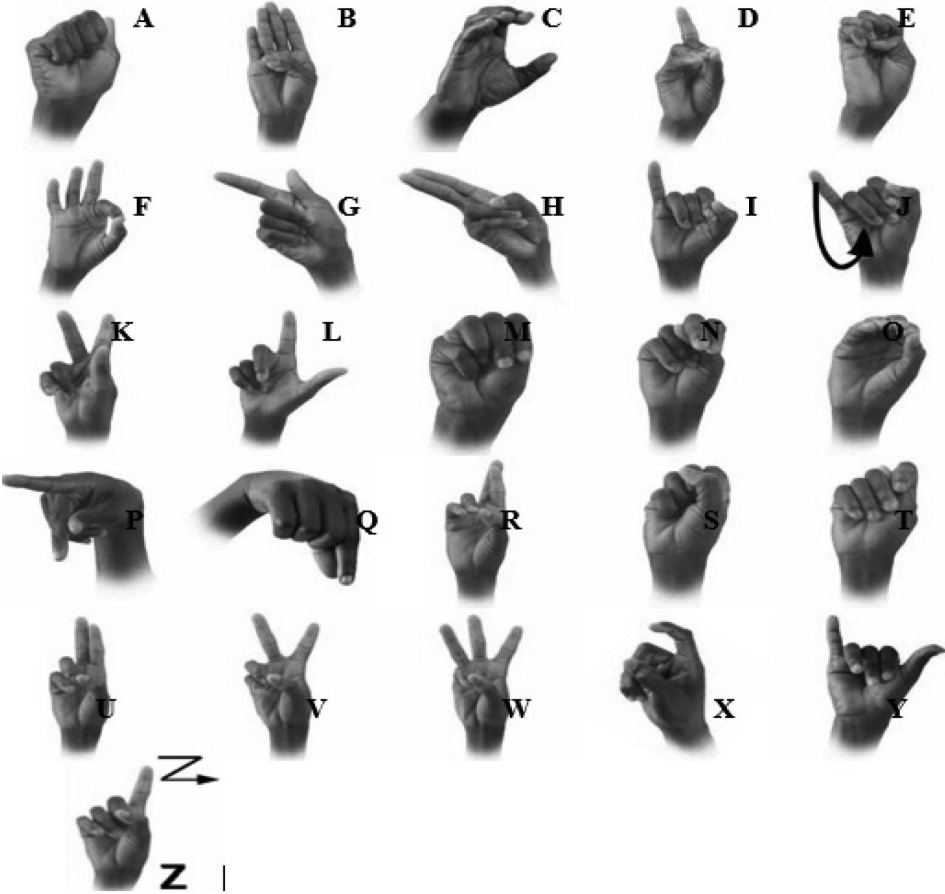
This chapter presented the results of iconic strategies used by signers and gesturers in Ghana. It discussed the preference for either instrument or handling by both signer and gesturers. Although both signers and gesturers used instrument and handling strategies, signers had a greater preference for instrument and gesturers a greater preference for handling except rural gesturers who demonstrated a high preference for instrument almost comparable to their preference for handling. Both signers and gesturers demonstrated a recurrent use of either instrument or handling iconic strategies across the semantic category of tools. The comparison between signers and gesturers demonstrate one facet of the emergence of a new sign language. The preference for handling forms becomes lexicalised as instrument when the language develops. Another finding of this research is the high use of instrument by rural gesturers (of Adamorobe), which seems to be related to AdaSL in the community. A study on the iconic patterns used by signers will further lead “to uncover mechanisms of representation and symbolic organization in gesture and language” (Padden et al. 2015:93).

The data from the signers were analysed for consistency and agreement, and AdaSL, a rural sign language had higher consistency and agreement in use of the same strategy for individual items than GSL, an urban sign language. It is also worth noting that lexical variants exist in GSL. Although lexical variants were not thoroughly investigated (because all the signers recruited for the study lived in proximity), when other lexical items in other semantic domains in three different dictionaries of GSL (published at different times) were compared, few lexical variants were identified. Further investigations with some native signers of GSL confirmed that some dictionaries have older versions of the signs and few variants are new entrants. For example, *mobile phone* signed with focus on the pole was used by 2 signers (deaf couple) who confirmed that the sign *mobile phone* with the Y-HS refers to TELEPHONE. Two recent dictionaries of GSL published in 2020 (online) and 2015 do not have the sign MOBILE PHONE but rather TELEPHONE depicted with the Y-HS as seen in figure 11. Irrespective of the existence of variants in GSL, there existed similarities between GSL and AdaSL (sharing the same preferences) and their preference for instrument strategy reiterate the fact that preference for instrument forms seems to emerge quickly in a new sign language (e.g. GSL).

The responses from the signers showed that the relationship between the linguistic form (phonology) and the meaning (semantics) involves series of associations that are

mapped from the phonological parameters and the meanings. The preference for a particular handshape is triggered by signers' profile and the construals of the object. Finally, there were handshapes that elicited fewer responses and others that elicited more responses in one-to-one mapping and many-to-one mappings respectively. The handshapes with higher numbers of association and frequency had lower semantic correlations because of the different associations linked to the handshapes. On the other hand, the handshapes with lower associations and frequency had stronger semantic associations for the handshapes. Iconic patterns among signers and gesturers reveals a mechanism for the emergence and development of sign languages and gestures.

APPENDIX. GSL manual alphabet (McGuire & Deutsch 2015)



References

- Brentari, Diane, Alessio Renzo, Jonathan Keane & Virginia Volterra. 2015. "Cognitive, cultural, and linguistic sources of a handshape distinction expressing agentivity". *Topics in cognitive science* 7(1), 95-123.
- De Vos, Roland Pau. 2015. "Sign language typology: the contribution of rural sign languages". *Annual Review Linguist* 1(1), 265-288.
- Dingemanse, Mark, Damián Blasi, Gary Lupyan, Morten Christiansen & Padraic Monaghan. 2015. "Arbitrariness, iconicity, and systematicity in language". *Trends in cognitive sciences* 19(10), 603-615.
- Ebling, Sarah, Konrad Reiner, Penny Braem Boyes & Gabriele Langer. 2015. "Factors to Consider When Making Lexical Comparisons of Sign Languages: Notes from an Ongoing Comparison of German Sign Language and Swiss German Sign Language". *Sign Language Studies* 16(1), 30-56.
- Edward, Mary. 2015. *Signing out: Linguistic contact and possible endangerment of the Adamorobe Sign Language*. Paper presented at the BAAL language in Africa SIG annual meeting on 22nd May 2015 at Aston University, Birmingham.
- Edward, Mary. forthcoming. *Iconicity as a pervasive force in language: Evidence from Ghanaian Sign Language and Adamorobe Sign Language*. Brighton: University of Brighton, Doctoral College.
- Evans, Vyvyan & Melanie Green. 2006. *Cognitive linguistics: An introduction*. Edinburgh: Edinburgh University Press.
- Frishberg, Nancy. 1987. "Ghanaian sign language". *Gallaudet encyclopedia of deaf people and deafness*, 778-779.
- Hou, Lina. 2018. "Iconic Patterns in San Juan Quiahije Chatino Sign Language". *Sign Language Studies* 18 (4), 570-611.
- Hwang, So-One et al. 2017. "Of the body and the hands: patterned iconicity for semantic categories". *Language and Cognition* 9(4), 573-602.
- Kimmelman, Vadim, Anna Klezovich & George Moroz. 2018. "IPSL: A database of iconicity patterns in sign languages. Creation and use". In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC-2018)*. Miyazaki, Japan: ELRA.
- Kiyaga, Nassozi & Donald Moores. 2003. "Deafness in Sub-Saharan Africa". *American annals of the deaf* 148(1), 18-24.
- Langacker, Roland. 2008. *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.
- Mandel, Mark. 1977. "Iconic devices in American sign language". In: Lynn A. Friedman (ed.). *On the other hand: New perspectives on American Sign Language*. New York: Academic Press, 57-107.
- McGuire, Caithlin & Casey Deutsch. 2015. *Ghanaian Sign Language Dictionary*. Morrisville: Lulu.com.
- Meir, Irit, Carol Padden, Mark Aronoff & Wendy Sandler. 2013. "Competing iconicities in the structure of languages". *Cognitive Linguistics* 24(2), 302-343.
- Miles, Mark. 2004. "Locating deaf people, gesture and sign in African histories, 1450s–1950s". *Disability & Society* 19(5), 531-545.
- Nyst, Victoria. 2007. *A Descriptive analysis of the Adamorobe Sign Language (Ghana)*. Amsterdam: Doctoral Dissertation, University of Amsterdam, Faculty of Humanities.
- Nyst, Victoria. 2016. "Size and shape depictions in the manual modality: A taxonomy of iconic devices in Adamorobe Sign Language". *Semiotica* 210, 75-104.
- Occhino, Corrine. 2016. *A cognitive approach to phonology: Evidence from signed languages*. Doctoral dissertation, The University of New Mexico, USA.
- Okyere, Alexander & Mary Addo. 1994. "Deaf Culture in Ghana". In: Carol J. Erting, Robert C. Johnson, Dorothy L. Smith & Bruce D. Snider (eds.). *The Deaf Way. Perspectives from the International Conference on the Deaf Culture*. Washington: Gallaudet University Press.
- Ortega, Gerardo & Asli Özyürek. 2016. *Generalisable patterns of gesture distinguish semantic categories in communication without language: Evidence from pantomime*. Talk presented at the 7th Conference of the International Society for Gesture Studies (ISGS7).
- Padden, Carol, Irit Meir, So-One Hwang, Ryan Lepic, Sharon Seegers & Tory Sampson. 2013. "Patterned iconicity in sign language lexicons". *Gesture* 13(3), 287-308.

- Padden, Carol, So-One Hwang, Ryan Lopic & Sharon Seegers. 2015. "Tools for language: Patterned iconicity in sign language nouns and verbs". *Topics in cognitive science* 7(1), 81-94.
- Perniss, Pamela. 2007. *Space and iconicity in German sign language (DGS)*. Nijmegen: Doctoral dissertation, Radboud University Nijmegen Nijmegen.
- Perniss, Pamela, Robin Thompson & Gabriella Vigliocco. 2010. "Iconicity as a general property of language: evidence from spoken and signed languages". *Frontiers in psychology* 1, 227.
- Sandler, Wendy & Diane Lillo-Martin. 2006. *Sign language and linguistic universals*. Cambridge: Cambridge University Press.
- Simone, Raffaele (ed.). 1995. *Iconicity in language*. Amsterdam: John Benjamins Publishing.
- Taub, Sarah. 2001. *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge: Cambridge University Press.
- Valli, Clayton, Ciel Lucas, Kristin Mulrooney & Miako Villanueva. 2011. *Linguistics of American sign language: An introduction*. Washington: Gallaudet University Press.
- Wilcox, Phyllis Perrin. 2000. *Metaphor in American sign language*. Washington: Gallaudet University Press.
- Wilcox, Sherman. 2002. "The iconic mapping of space and time in signed languages". *Advances in consciousness research* 41, 255-282.
- Wilcox, Sherman. 2004. "Cognitive iconicity: Conceptual spaces, meaning, and gesture in signed language". *Cognitive Linguistics* 15(2), 119-148.
- Wittenburg, Peter, Hennie Brugman, Albert Russel, Alex Klassmann & Han Sloetjes. 2006. "ELAN: a Professional Framework for Multimodality Research". *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*.

Part IV



Social and communicative aspects of languages

Phraseological units involving body-part terms: a corpus based analysis of Hausa to English translation

Abstract

The article focuses on the comparison of structures involving body-part terms in one language and their equivalents in another. The analysis is based on examples extracted from three novels published in both Hausa and English versions. Ten items representing the terms for body parts (mouth, stomach, head, back, front, hand/arm, eye, ear(s), leg/foot, heart) were checked in the parallel texts and compared in terms of their lexical equivalency. The analysis revealed important differences between the two languages in the use of body-part terms. A significantly higher number of body-part terms in the Hausa text than in its English counterpart is motivated by structural properties of the Hausa language in which these terms are involved in coding grammatical relations; it also has reference to the patterns of conceptualization in which body-part terms have various semantic extensions and metaphoric functions.

Keywords: phraseology, body-part terms, grammaticalization, cultural metaphors, Hausa

1. Introduction

Phraseology is defined as the co-occurrence of a form or a lemma of a lexical item and one or more additional linguistic elements of various kinds which function as one semantic unit in a clause or sentence and whose frequency of co-occurrence is larger than expected on the basis of chance (Gries 2008: 6). This definition opens a variety of interpretations of what phraseological units (hence PU's) are and what information can be extracted from the analysis of word combinations that cannot be decoded by the word alone.

A phraseological unit is therefore viewed as a stable, cohesive combination of words with a fully or partially figurative meaning (Naciscione 2010: 32). PU's do not exceed the boundaries of one sentence, in which they are semantically and structurally more cohesive than words, but in discourse use they are flexible and involved in a series of developments, acquiring additional stylistic features. Language specific aspects of phraseologisms create many problems in the study of language, which are also manifested in translation and/or second language teaching.

Along with grammatical, figurative or stylistic functions, phraseological units are known to exploit some parts of lexicon to create new meanings, different from their source meanings. Body part terms represent the group of words that are widely used in many (if not all) languages of the world in different types of metaphors, in which they create various conceptualization patterns (Brenzinger & Kraska-Szlenk 2014).

Scholars have extensively discussed body part terms from various perspectives. Submissions having a comparative perspective such as (Johnson 1987; Sharifian et al. 2008; Maalej & Yu 2011; Kraska-Szlenk 2014) or specifically related to Hausa (Pawlak 2014; Almajir 2013, etc.) confirm that body-part terms (Bpt's) in various languages show many similarities in ways of extending their meanings into abstract domains. On the other hand, structural changes in the process of their grammaticalization and various conceptualization patterns create differences in their function between languages (Heine 2014). Stylistic values and association with language registers may additionally differentiate the use of body-part terms in particular languages. It seems interesting to look at the semantics of body-part constructions from the perspective of translation.

2. Aim and scope

The present chapter focuses on the use of body-part terms in phraseological units and gives the analysis a comparative perspective. By using parallel texts from two different languages, the structural and semantic properties of Bpt's as well as their cultural context are recognizable.

In particular, the comparison of structures involving body-part terms in one language and their equivalents in another, shows the differences between languages in the conceptual patterns and grammaticalization, which is widely certified for this part of lexicon. Therefore, I am trying to analyze the function and use of body-part terms in contemporary Hausa literary writings which have their translated versions in English. These are:¹

1. *Amina* by Mohammed Umar (English 2005; Hausa 2011)
2. *Masoyan zamani/ Nemesis* by Ado Ahmad Gidan Dabino (Hausa 1993; English 1995)
3. *In da So da Kauna/ The Soul of my Heart* by Ado Ahmad Gidan Dabino (Hausa 1991; English 1993).

For further references, the books are coded as follows:

1. *Amina* [Hausa] (AmnH) / *Amina* [English] (AmnE)
2. *Masoyan zamani* (MszZ) / *Nemesis* (Nmss)
3. *In da So da Kauna* (IndS) / *The Soul of my Heart* (SomH).

The idea here, as stated above, deals with the question of how the Hausa PU's are translated into English and whether there is an equivalency of the use of Bpt's in the two languages. With reference to Heine (2014:1) stating that the relationship between a body part term and conceptualization has a varying degree of possibilities, the idea is based on assumption that Bpt's as lexical items are not used in the same way in different languages.

The two languages for this study, namely Hausa and English, are genetically unrelated and geographically distant. They are classified into two distinct language families, Afro-Asiatic and Indo-European respectively. They have had separate ways in their development, but they have many opportunities to meet in modern times, There is also a significant number of translation works conducted between them, which include literary writings and other texts. Of course, greater percentage of the translations is from other sources to

¹ The first book appeared in the English version first.

Hausa, though efforts are now being made to translate Hausa into other languages. Native scholars have identified some problems associated with translations from and to local languages. Referring to translations from some Nigerian languages into other languages, like English, it was pointed out that they have often focused on lexical translation, using translation strategies such as semantic shifts and transliteration (Asobele 1999). However, translation strategies certainly go beyond the word level to embrace the phrase, sentence and discourse and this also applies to translation of texts from Nigerian languages into English (Azare 2015; Azare 2019). The comparison of Hausa and English texts is misstated by these external factors, but the fact of creating original text once in English, once in Hausa makes the two parallel texts close equivalents.

The comparison of Hausa PU's involving Bpt's with their equivalents in English proposed by translators allows discussing various dimensions of using Bpt's in phraseologisms. The following aspects are the subject of a more detailed analysis: the function (lexical or grammatical) of body-part terms in Hausa in relation to English, the use of Bpt's in conceptual metaphors and the frequency of usage of the selected PUs in the two languages.

3. Body-part terms as a basis for comparative works

Body-part terms are lexical items which have a grammatical status of nouns. When used independently, their meaning refers to the parts of the body and they may have either singular or plural form (e.g. sg. Hau. *hannu*, Eng. *hand*, pl. Hau. *hannaye*, Eng. *hands*). When used in phraseological units, only one form is admitted, usually it is a single noun, e.g. Ha. *hannun riga* 'sleeve of a shirt' (lit. hand of a shirt), Eng. *hand-bag* 'woman's bag for money, keys, handkerchief, etc.; *ya shiga hannunmu* 'it is in our possession' (lit. he entered our hands), Eng. *it is in our hands* 'we look after it, we manage it'. As the above examples indicate, the units involving body-parts may share some properties in coding the meaning in different languages, but they are language specific.

The Bpt's used in this study have been extracted from the text of Umar's and Gidan Dabino works used as a corpus. Ten items representing the terms for body parts (mouth, stomach, head, back, front, hand/arm, eye, ear(s), leg/foot, heart) were checked in both Hausa and English parallel texts and compared in terms of their lexical equivalency², i.e. whether the Hausa term has its counterpart in the English text or not. On the other hand, the Hausa equivalents of the English terms were also examined. Since the analysis focused on phraseological units involving Bpt's, a number of occurrences have been omitted as irrelevant for the study, and finally the corpus was made up of 410 PU's sipped from over 7000³ occurrences. The collected material serves as a basis for comparison between the two languages.

In the course of our findings, we discovered important differences between the two languages in the use of body-part terms. In the analysis, both quantitative and qualitative

² The Hausa terms for *hannu* 'haand, arm' and *kafa* 'leg, foot' have two lexical equivalents in English, following the differences in coding the meaning between the two languages. Similarly the noun *kunne* 'ear', which may be contextually interpreted as 'ears'.

³ The search was automatically done, therefore, the record for *baya* may include such words as *bayani*, but these examples are rare and insignificant with regards to frequency.

criteria were taken into account. The examples were checked along translation lines of frequency, and then the conclusions on the use of Bpt's in grammatical function and in creating conceptual metaphors were made. Some cultural nuances were also drawn from the interpretation of data.

4. Body-part terms in Hausa phraseology

Hausa body-part terms are used in various types of structures in which they function either as part of grammatical patterns or as a lexical element of phrasal expressions or part of other stable combinations of words. For the purpose of the present analysis, the following structures have been identified:

- a. Body-part terms used as grammatical elements
- b. Body-part terms used as nominal complements of phrasal verbs
- c. Body-part terms as cultural metaphors

4.1. Body-part terms as grammatical elements

A large part of body-part terms in Hausa has been involved in the process of grammaticalization, which has changed them to grammatical markers. Among the so-called genitive prepositions (Newman 2000: 470), many are built on body Bpt's which are used to express spatial relations. The most commonly used prepositions of this kind include: *kai/kan* 'on' (lit. head-of), *baki-n* 'edge of, (lit. mouth-of), *baya-n* 'behind' (lit. back-of), *ciki-n* 'inside' (lit. stomach-of), *gaba-n* 'front' (lit. front/breast-of)⁴. It means that these body part terms, often used in constructions, over time, become grammatical elements indicating location. Heine (2014: 17) views them as grammaticalization denotators, such as *baki* used in *...sai ta zauna a bakin gado tana ba danta mama*" (AmnH 114). The term *baki* which refers literally to *mouth* is no longer translated as such and the phrase as above has the meaning '...and she sat **at the edge of the bed** breastfeeding her child'. Similarly, in the other examples like *...na bakin kogi* (IndS:79) the word *baki* refers to 'edge of the river or **river bank**'. Other examples are related to other Bpt's which are seen as grammatical markers in Hausa constructions, like *baya*, the term literally referring to 'back', but in the phrase *Daga nan suka shiga bayan motar* (IndS: 91) 'from there they entered **the trunk of the car**' it is spatial marker, as well as *gaba* 'front/breast' in *suka tsaya a gaban mai shari'a* (MsyZ: 71) 'they stood **before the judge**', *ciki* 'stomach' in *ba irinta cikin garin* (AmnH:5) '*it's very unique, you cannot find it in the town*' or *kai* 'head' in *kan tebur* '**on the table**'. This is to show how a large number of Bpt's function as grammatical elements in Hausa language and their status is not to be compared with the equivalents in English, in which body-part terms are not so clearly manifested in locative constructions. Some of these terms have also evolved as temporal or reflexive markers and it makes the occurrence of some body part terms in the Hausa texts relatively high (cf. section 5). In English translation, neither the phrase *bayan sun zo* (or *bayan zuwansu*) '**after** they came' nor *shi da kansa* '**he himself**' contain any body-part term evolved to the function of a grammatical marker.

⁴ As in all regular compounds, genitive prepositions have a genitive linker (mostly -n, rarely -r).

Other body-part terms as suggested in this study may not be so highly grammaticalized as the ones exemplified above. However, on a large scale they are used as constituents of phrasal verbs or compounds.

4.2. Body-part terms in phrasal verbs and nominal compounds

Another aspect in this study that allows investigation into the phenomenon of phraseological units involving body-part terms is the occurrence of phrasal verbs which are fixed verb-object collocations⁵. Many of them involve Bpt's which follow verbs such as *sa* 'put', *ci* 'eat', *ja* 'pull', *sha* 'drink', *juya* 'turn', *hada* 'join', *ba (da)* 'give', etc., e.g. *sa baki* 'interfere' (lit. put mouth), *ba da baki* 'coax' (lit. give mouth), *ja kunne* 'warn' (lit. pull ear), *ci gaba* 'proceed' (lit. eat front), *sha kansu* 'manage to get ahead (of them)' (lit. drink their head), *juye kansa* 'control (him)' (lit. turn his head), *hada baki* 'conspire' (lit. combine mouth), *hada kai* 'cooperate' (lit. join head) and many others. These phrasal categories with mono-syllabic or bi-syllabic verbs often utilize Bpt's to introduce some conceptualization models based on metaphoric or metonymic representation (cf. Jaggar & Buba 2009). They are found both in regular and less regular patterns. They represent language-specific patterns of conceptualization and their meaning is recognized beyond the word level. Along this line, we will investigate the English counterparts and see whether expressing the meaning also follows the pattern with body-part term. In many cases, however, the strategy of coding the meaning is different, e.g. when the phrase *ci gaba (da)* 'proceed, continue' (lit. eat front) is used, as in *Suka ci gaba da hirarsu*' (AmnH:1) translated as 'They **continued** their light-hearted chatter' (AmnE: 1).

Still under this category of idiomatic phrase, we have the nominal compound that constitutes one grammatical unit (i.e. it functions as NP of a sentence and cannot be discontinuous). This class also utilizes body-part terms in their specific meaning, though conceptualization is tight to the metaphoric or metonymic intentions. For instance: *farin-ciki* (lit. white-stomach) 'happiness', as in *kuma na yi matukar farin cikin ganinta* 'I was extremely **happy** seeing her'. others could be *goyon baya* (lit. something or baby carried on the back) 'support', *bakin-ciki* (lit. black-stomach) 'sadness', etc.

In the present analysis, the translated version of the phrasal verbs and nominal compounds will be compared with their original Hausa forms involving body-part terms. The question is whether this part of lexicon being a source of conceptualization patterns finds its equivalents in the English expressions.

4.3. Body-part terms as cultural metaphors

Languages differ significantly in the way they express and use what is in their culture. As a fundamental in every society, culture plays such a role of motivating meaning using various metaphors established within the language. Metaphors on the other hand, encapsulate meaning that cannot be literally deducted. Understanding metaphors always goes beyond the word's literal meaning (Gibbs et al. 1997). In this study, we have found that in Hausa cultural metaphors are used in phraseological units during and for discourse

⁵ Termed as idiomatic phrasal verbs in (Newman 2000: 260).

purposes. The meaning, as in most cases, cannot be seen at par with what can be arrived in another language. One of the striking features of the Hausa language in this regard, is the extensive use of body-part terms in making these cultural metaphors and it may be regarded as a distinctive feature or an orientation point of the language. Preliminary analysis showed that the frequency of occurrence in this aspect is by far different from what can be obtained in other languages (English in this case). The PU's in this class can take the form of figurative expressions as metaphors, proverbs, culture key words etc. When used in their literal meaning, body-part terms refer to the animals or birds which are heroes of story tales or have a specific cultural connotation, for instance *bakin kada* (lit. crocodile's mouth) in *sun ceto shi daga bakin kada* (AmnH:24) 'they saved him **from danger**'. We can see that *bakin kada* is used to indicate a genuinely dangerous situation. However, literal look at the phrase will indicate the nominal function of *baki* and the reference to actual animal crocodile. The cultural motivation here is neither on the *baki* nor on the animal crocodile. The information has its cultural connotation by using a metaphoric expression. In another expression, ...*ana sheka ruwan sama kamar da bakin kwarya* (MsyZ:64), (lit. 'up to the edge of the calabash'), the body-part term is used as a spatial marker, but with special reference to the traditional pot. In the English translation ('it's raining cats and dogs'), however, coding the meaning is connected with the use of the other key-words.

The items which function as grammatical elements increase the occurrence of body part terms in phrasal expressions, but also some other body part terms are used in various types of phraseologisms, such as *ido* 'eye' in *duk yarinyar da aka gan ta tare da su ba ta da sauran mutunci a idanun jama'a* (MsyZ:32) meaning 'any girl seen with them has lost her public respect' (lit. **eyes of the people**). Similarly, *kafa* 'foot' in *ta zamar mana karfen kafa* (MsyZ:14) meaning 'she has become a burden (lit. **iron foot**) to us'. In Hausa to English translation, they have been rendered differently due to language peculiarities and further analysis shows that they affect the degree of equivalence in the PUs between the two languages.

One of the objectives here is to find out the extent to which body part terms are used in Hausa language as compared to English. The issues of cultural nuances and language orientation patterns in using Bpt's would be investigated.

5. Body-part terms in translations

The comparison of the Hausa phraseologisms involving Bpt's and their English equivalents in the novels allows distinguishing the following levels of their adequacy:

- a) phrasal expression with Bpt's in Hausa rendered with simple lexical equivalent (without Bpt) in English
- b) phrasal expression with Bpt's in Hausa rendered with lexical (phrasal) equivalents involving Bpt in English
- c) metaphoric expressions with Bpt's in Hausa rendered with established phrases in English
- d) proverbs rendered with proverbs.

Following the distinguished classes, some examples of phraseological units involving body-part terms in Hausa and their English equivalents have been extracted from the sources.

5.1. Bpt rendered with zero Bpt

Most phrases which are regular phrasal verbs have their direct equivalents in English terms which are single verbs, e. g.:

Hausa phrase: ...*Uwargida ta sa baki* (AmnH: 44) (lit. let [senior] wife **put mouth**)
 English equivalent: 'she appealed to Talatu to **intervene**'

Hausa phrase: *Aka ciyar da ni gaba* (AmnH:5) (lit. I was made to **eat front**)
 English equivalent: 'I was **promoted**'

Hausa phrase: *Sukan hada baki* (AmnH: 65) (lit. they usually **join mouth**)
 English equivalent: 'they actively **collaborate**'

Hausa phrase: *Duniya ta juya mata baya* (AmnH:45) (lit. the world turned its back to her)
 English equivalent: 'She became **alienated**'

The Hausa nominal compounds involving body-part terms represent language-specific patterns of conceptualization (*farin ciki* 'happiness', lit. white stomach) which are not copied in the English equivalents, e.g.:

Hausa phrase: *Cikin koshin lafiya da farin-ciki* (AmnH:24) (lit. in full health and **white stomach**)
 English equivalent: 'healthier and **happier** lives'

Similarly, expressing emotions through body-part terms is based on different conceptual patterns in the two languages, e.g.:

Hausa phrase: *Zuciyar Amina ta tafi yawo* (AmnH: 159) (lit. Amina's **heart** went on an errand).
 English equivalent: 'Amina lost **interest**'

The examples as above confirm that in the parallel texts, the frequency of body-part terms is supposed to be higher in Hausa texts than in English texts.

5.2. Bpt rendered with any or another Bpt

In some groups of expressions, the body-part orientation is kept in the two language versions, as, for example, when the term *baya* 'back' is used to code spatial and time orientation or to describe emotions though *zuciya* 'heart':

Hausa phrase: *Ra'ayin mayar da hannun agogo baya* (AmnH:35) (lit. returning hands of the clock **backwards**)
 English equivalent: 'propagate **backward** ideas'

Hausa phrase: *Tuna baya shine roko* (AmnH:74) (lit. recalling **earlier days** is a request)
 English equivalent: 'she recalled **earlier days**'

Hausa phrase: *Zuciyarta ta yi fari kwal* (AmnH:236) (lit. **her heart** became snow white).
 English equivalent: 'her **heart** lifted'

5.3. Cultural metaphors rendered with established statements

The differences of cultural code between Hausa and English are manifested in PU's which refer their understanding to cultural realia, as in the following examples:

Hausa phrase: *Suka ceto shi daga bakin kada* (AmnH:24) (lit. they saved him from the **crocodile mouth**)
 English equivalent: 'rescued him in a **daring operation**'

Hausa phrase: *Mata masu hannu da shuni* (AmnH:2) (lit. Women with **indigo in their hands**)
 English equivalent: 'the **richest** women'

5.4. Proverbs rendered with proverbs

Translation of proverbs is motivated by rendering the sense rather than translating the phrasal constituents. The use of body-part terms in phrasal expressions is motivated by cultural connotations of culture key-words rather than by the terms themselves.

Hausa phrase: *An fede biri har wutsiya* (AmnH:200) (lit. flay the skin of **monkey** to the tail)
 English equivalent: 'strip the **chicken** of its feathers' (i.e. divulge all)

6. Body-part terms in frequency of occurrences

Checks on the three parallel texts in Hausa and English respectively were made as a corpus search in which the list of body-part terms was established on the basis of the most common Hausa forms. The comparison includes 10 Bpt's selected for this study and search was done automatically on Hausa forms in the Hausa texts on the one hand, and on their English equivalents in the English texts on the other. For instance, *baki* was searched in English as '**mouth**' and the corresponding frequency was entered accordingly.

The number of occurrences in each of the three books has been summarized in the table below:

The corpus search revealed that, the selected Bpt's appear more frequently in the Hausa texts than in the English texts or – respecting the source of comparative analysis – in English rendition the Bpt's do not appear within a similar range as in Hausa text. This is clear not only from individual novels, but from the overall corpus considered for this analysis. Therefore, we envisage Hausa having a higher frequency of Bpt's than English and this is due to high involvement of Hausa body-part terms in phraseology.

S/N	Key words	<i>Amina</i> (Hausa 371 pages)	<i>Amina</i> (English 252 Pages)	<i>Masoyan Zamani</i> (Hausa 76 pages)	<i>Nemesis</i> (English 110 pages)	<i>In da So</i> (Hausa 304 pages)	<i>Soul of my Heart</i> (English 100 ⁶ pages)	Total of Each Hausa/English	
1.	Baki- (n) Mouth	81	10	36	01	125	03	242/14	
2.	Ciki- (n) stomach	662	02	244	03	483	00	1389/05	
3.	Kai/Kan Head	383/986	60	243/88	06	391/589	06	2680/72	
4.	Baya- (n) Back	474	95	213	23	409	11	1096/129	
5.	Gaba- (n) front	456	34	82	01	172	00	710/35	
6.	Hannu-(n) Hand/arm	157	37/02	25	18/1	140	26/03	322/87	
7.	Ido-(n) Eye(s)	96	71	15	09	154	05	265/85	
8.	Kunne Ears	11	07	05	01	15	01	31/09	
9.	kafa Legs/foot/feet	34	09/14	08	00	30	02/0/0	72/25	
10.	Zuciya/ zuci Heart/mind	74	16/24	51	16/21	111	15/09	236/101	
		Total 7,045/558							
		92,1%/ 7.9% = 100%							

Table 1. Body-part terms frequency of occurrences in Hausa and English parallel texts

The frequency of particular items indicates that some body-part terms are more readily used than others. Key words 1-5 seem to have higher frequency of occurrence when compared with the rest. Thus, *baki*, *ciki*, *kai*, *baya* and *gaba* have taken the toll with regard to frequency. As earlier stated and explained elsewhere in this paper, their high frequency is directly connected with their involvement in the process of grammaticalization. Other Bpt's that are less involved in the process of grammaticalization appear to have less frequency. Thus, *hannu*, *ido*, *kunne*, *kafa* and *zuciya* are also used in expressions which extend their meaning of body-part terms to other notions, but they are rarely used in grammatical function.

⁶ Our checks on this book reveals that a substantial number of pages have not been translated, hence the low number of pages in the English version.

With reference to the 10 Bpts, total score of **558** occurrences in English to **7,045** occurrences in Hausa is a clear indication of a higher frequency of body-part terms usage in Hausa than in English. The percentage is by far incomparable (Hausa examples represent a significant proportion of the whole sum, English equivalents reach only **7.9%** of the whole sum). It shows that body-part terms in Hausa and English function differently, both at the level of grammar and as a source domain of conceptualization. The analyzed examples indicate the tendency of substituting a PU involving Bpt with an established statement (usually without a Bpt) or with just a simple verb in the English language. Other possibilities could be rendering proverbs with proverbs which might or might not involve Bpt in the English equivalent. As a result, only few PU equivalents may present any or another body part in the translation.

7. Summary and concluding remarks

The analysis of the use of body-part terms in English and Hausa is directed at comparison of languages which are geographically distant and culturally different. The analysis of the use of lexicon, which is considered universal, shows the main differences between languages in the status of these terms and their use as a source domain of various conceptualizations. The study discovered that culturally determined conceptualization of ideas in the Hausa language is manifested through the extensive number of body-part terms, therefore the conceptualization is often based on human body as a point of orientation. The same cannot be said for English, as some terms are created irrespective of these motivations.

We have also discovered that the grammatical patterns of the two languages employ different schemas to achieve the intended rendition. In the Hausa examples of PU's, both regular and less regular phrasal verbs (involving Bpt's) and or compounds were found in the texts. The results showed that most of those structures are realised with just a single verb or noun when rendered into English. This is a clear manifestation of a deeply rooted use of body-part terms in Hausa.

A very rare phenomenon was also established that a Bpt in Hausa is substituted in English with another Bpt. This indicates that both languages operate some common traits of utilizing Bpt's in apportioning meaning to certain items. However, the frequency proves rather low.

This study discovered that body-part terms in Hausa are often used in expressions which function as cultural metaphors. As common language forms that are readily available in the language they code many other lexical meanings and express grammatical relations. Likewise, they are also used in proverbs which are integral part of Hausa culture and play a vital role in communication.

The two parallel text analyzed here put stress on differences rather than on similarities between languages in the use of body-part terms. The stylistic and communication aspects of the phraseology involving Bpt's needs further research.

References

- Almajir, Tijjani Shehu. 2013. "The Polysemy of Body Part Terms in Hausa within the Frame of Image Schemas". *Studies of the Department of African languages and Cultures* 47, 91-109.
- Asobele, S. J. Timothy. 1999. *New Perspective in the Training of Tranalators and Interpreters in Nigeria*. Lagos: Print View Publication.
- Azare, Yakubu Magaji. 2015. "Creativity and Lexical Innovation in Hausa – English Renderings". In: Gian Claudio Batic & Sergio Baldi (eds.). *Selected Proceedings of the Symposium on West African Languages (Naples, 27-28 March 2014)*. *Studi Africanistici* [Serie Ciado-Sudanese, Vol. 7]. Naples: Il Torcoliere Università L'Orientale, 309-320.
- Azare, Yakubu Magaji. 2019. "Phraseological units of Shakespeare's "Merchant of Venice" in the process of translation into Hausa". In: Nina Pawlak (ed.) *Linguistic evidence of cultural distance. Hausa in cross-cultural communication*. Warszawa: Elipsa, 99-111.
- Brenzinger, Mathias & Iwona Kraska-Szlenk (eds.). 2014. *The Body in Language: Comparative Studies of Linguistic Embodiment*. Leiden: Brill.
- Gibbs, Raymond W., Josephine M. Bogdanovich, Jeffrey R. Sykes, & Dale J. Barr. 1997. "Metaphor in Idiom Comprehension". *Journal of Memory and language* 37 (2), 141–154.
- Gries, Stefan Th. 2008. "Phraseology and linguistic theory: a brief survey". In: Sylviane Granger & Fanny Meunier (eds.). *Phraseology: An interdisciplinary perspective*. Amsterdam: John Benjamins, 3-26.
- Heine, Bernd. 2014. "The Body in Language: Observation from Grammaticalisation". In: Matthias Brenzinger & Iwona Kraska-Szlenk. *The body in Language: Comparative Studies of Linguistic Embodiment*. Leiden: Brill, 13-32.
- Johnson, Mark. 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination and Reason*. Chicago and London: The University of Chicago Press.
- Jaggar, Philip & Malami Buba. 2009. "Metaphorical extensions of 'eat' ⇒ [overcome] and 'drink' ⇒ [undergo] in Hausa". In: John Newman (ed.). *The Linguistics of Eating and Drinking*. Amsterdam and Philadelphia: John Benjamins, 229-251.
- Kraska-Szlenk, Iwona. 2014. "Extending Body-Part Terms in the Domain of Emotions". In: Matthias Brenzinger & Iwona Kraska-Szlenk. *The body in Language: Comparative Studies of Linguistic Embodiment*. Leiden: Brill, 52-70.
- Maalej, Zouheir & Ning Yu (eds.). 2011. *Embodiment via Body Parts: Studies from Various Languages and Cultures*. Amsterdam/Philadelphia: John Benjamins.
- Naciscione, Anita. 2010. *Stylistic Use of Phraseological Units in Discourse*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Newman, Paul. 2000. *The Hausa Encyclopedic Reference Grammar*. New Haven: Yale University Press.
- Pawlak, Nina. 2014. "Notion of SELF in Hausa". In: Matthias Brenzinger & Iwona Kraska-Szlenk (eds.). *The body in Language: Comparative Studies of Linguistic Embodiment*. Leiden: Brill, 140-160.
- Sharifian, Farzad, Rene Driven, Ning Yu & Susanna Niemeier (eds.). 2008. *Culture, Body, and Language: Conceptualizations of Internal Body Organs across Cultures and Languages*. Berlin/New York: Mouton de Gruyter.

Sources

- Umar, Mohammed Kabir. 2005. *Amina* (English version). Africa World Press.
- Umar, Mohammed Kabir. 2014. *Amina* (Hausa version). Africa World Press.
- Gidan Dabino, Ado Ahmad. 1993. *Nemesis*. Gidan Dabino Publishers Enterprises.
- Gidan Dabino, Ado Ahmad. 1993. *Masoyan zamani*. Gidan Dabino Publishers Enterprises.
- Gidan Dabino, Ado Ahmad. 1991. *In da So da Kauna*. Kano, Nigeria: Nuruddeen Publication.
- Gidan Dabino, Ado Ahmad. 1993. *The Soul of my Heart*. Gidan Dabino Publishers Enterprises.

Linguistic repertoires of refugees in Internally Displaced Peoples' (IDP) camps in North East Nigeria

Abstract

Violent crises orchestrated by the activities of an Islamist extremist group in Nigeria, Boko Haram, besides resulting in the destruction of lives and properties in many communities in North East, Nigeria, contribute significantly to mobility of languages. A vast number of survivors who fled from the volatile areas inadvertently moved to different refugee camps designated by the Nigerian government as Internally Displaced Persons (IDP) camps. Given the fact that the linguistic ecology of North East Nigeria is extremely diverse, the situation where people from diverse linguistic backgrounds co-habit in camps under austere conditions would yield to some sociolinguistic issues including language politics and language endangerment, among others. The IDP camps attract the attention of international agencies such as UNHCR and national groups who send their representatives for sundry purposes. The language of the visitors is often a global language, English, whereas the displaced people in the camps are from diverse linguistic backgrounds with Hausa as the most popular language. In this situation, some persons in the camp are bound to encounter some language and communication challenges. Using some principles from Ecolinguistics, this study examined the linguistic repertoires of people living in some selected IDP camps in North East Nigeria. The instrument for data collection were interviews, observation and focus group discussions. The information from the study highlights the challenges of the traumatized persons who are grappling with communication challenges, the diverse sociolinguistic issues derived from the linguistic practices and the precarious situation of the languages on the move. The information from the study directs attention to the current trends in the bustling linguistic ecology of Nigeria and the need to take proactive measures to maintain the rich cultural heritage.

Keywords: Linguistic repertoires, violence, language mobility, Internally Displaced Persons, language starvation, language endangerment

1. Introduction

The bustling West African linguistic ecology has been vigorously examined and described in scholarship (Ballard 1971; Shimizu 1980; Crozier & Blench 1992, Elugbe 2009b, Blench 2012). The plural linguistic situation in Nigeria is often referred to as a typical example of West African linguistic diversity (Spencer 1971). Hence, Nigeria's linguistic ecology has been described with superlative terms such as – extremely linguistically diverse, crassly diverse (Elugbe 2009a) hotbed of diversity and super diversity (Harnischfeger et al. 2014).

In the attempt to provide a vivid description of Nigeria's bustling linguistic heritage, some studies purposely or inadvertently reflect the distribution of languages in Nigeria across speech communities in such a way that the diverse Nigerian languages are often neatly tucked into geographical boundaries within the Nigerian space (Meek 1931; Ballard 1971; Shimizu 1980; Webster 1993; Blench 2012; Ethnologue 2018). The assumption is that Nigeria consists of different ethnic nationalities or language groups that are occupying easily identifiable speech communities and that indigenous languages in Nigeria are spoken in different specific geographical spaces within the country.

This narrative does not reflect the current situation. The reality is that there is hardly any homogenous speech community within the present-day Nigerian space. Several languages are in contact in towns and villages in Nigeria yielding to fading linguistic boundaries. The situation also gives rise to some communication and sociolinguistic issues which include language starvation, linguistic and identity fluidity, language politics, language shift, language dominance, language threat, language death, among others. However, in conversations especially in relation to governance, the account is usually that a particular language and the speakers can be located in specific geographical spaces. This type of postulation is usually sponsored for political reasons.

2. Ecolinguistics

Ecolinguistics is a paradigm in sociolinguistics research which incorporates the study of ecological context of languages. It accounts for ecological factors as well as its consequences on the languages under study. It attempts to establish a connection between linguistic and biological diversity (Muhlhausier 2003). Ecolinguistic paradigm has been employed in linguistic research to examine issues relating to mobility – globalisation as well as the current dynamics of movements in modern immigration contexts (Hatoss 2013). It is also employed in the study of linguistic diversity and the embedding of traditional knowledge in local languages, which is described as 'the story we live by' (Stibbe 2015). Ecolinguistics principles have been resourceful in the study of the language of migrants. Hatoss (2013) utilized Ecolinguistics approach to study the daily linguistic practices of refugees, the vibrancy of the languages in use and the ecological factors that influence the socio-economic participation of the refugees in the immediate community and in the society at large.

One key goal of Ecolinguistics research is to protect cultural diversity as well as the linguistic diversity that supports it (Harmon 1996; Nettle & Romaine 2000; Muhlhausier 2003; *Terralingua* 2008). It investigates the current linguistic ecology of communities to determine the vibrancy of the languages by examining the daily linguistic practices of the people in their contexts (Hatoss 2013). Information from a study of this nature can be utilized in language planning.

The approach employs many tools to investigate the trends in the linguistic ecology of communities. It adopts different methods in investigation – quantitative and qualitative tools as well as discourse analysis. Ecolinguistics can be applied in linguistic studies in two broad senses. It can be used to study the physical ecological system of a particular language. It can also be used in metaphorical sense in the study of linguistic ecology

of a language. Generally, it adopts a broad approach in its study of language as it considers the intricacies encapsulated in the nature of linguistic realities.

3. Language mobility in Nigeria

Language mobility can be explained as the movement of a group of speakers with their language from their natural habitat to other locations. Many factors contribute to the evolving flux in the linguistic situation as well as distribution of languages across geographical spaces in Nigeria. One key factor is increased mobility – the movement of people and their languages from one place to another. There appears to be constant and unpredicted movement of people within and outside Nigeria. Mobility within the Nigerian geographical boundaries can be seen from two dimensions – voluntary and forced mobility. In the case of voluntary mobility, a person, family or group can decide to move from one place to another, temporarily or permanently, due to sundry reasons. In the latter case, an individual, a family, clan, group, etc., is forced to move, in most cases, from the ancestral home to avert annihilation. Forced mobility is one key contributory factor to the current linguistic and sociolinguistic situations confronting the rich linguistic heritage of Nigeria.

Some salient factors contribute to increased mobility of languages in Nigeria:

- i. Dynamics of globalisation, which involves the collapse of cultural as well as linguistic boundaries. Globalisation has contributed to trends in linguistic practices and identities (Nassenstein & Hollington 2016).
- ii. Breakthrough in technology, which has made it possible for languages to move across digital spaces even if it is never accompanied with physical movement.
- iii. Tourism contributes to mobility of languages. Tourism has made it possible for several languages in different parts of the globe to come into contact with Nigerian languages. On the other hand, Nigerian languages have also moved with their speakers to several parts of the country and around the globe.
- iv. Religion – Christianity and Islamic religions – contribute to the mobility of people as well as their languages across spaces in Nigeria and outside Nigeria.
- v. Urbanisation is a key candidate in mobility of persons and languages. It contributes to the linguistic diversity situation in current Nigeria. Urbanisation in Nigeria is often characterised by unequal distribution of resources across the cities, states and zones. The concentration of resources and facilities such as unity schools, government medical facilities, etc., in particular places, especially urban areas yields to the influx of people to such places. This situation leads to myriad languages in contact in some key urban centres in Nigeria. The speakers of minor languages in urban centres often dwindle while community of speakers of the major languages often attract more speakers and continue to increase. Urbanisation also contributes to identity fluidity.

The factors identified above contribute to voluntary mobility to a large extent.

- vi. Agriculture contributes to forced movement of people, culture as well as languages. For example, agriculture, basically food production, is a major reason why clans

move from one location to another. Farmers in desperate search for more fertile farmlands or waters for fishing in Nigeria move from one location to another even across states. For example, Tiv speakers in search of fertile farmlands can be found in different states and regions in Nigeria. Thus, Tiv is recorded as one of the indigenous languages in their ancestral home Benue, in other states in North Central and North East Nigeria.

- vii. Natural disasters such as flood, famine, desert encroachment, especially disasters which affect food production are contributory factors to forced mobility in Nigeria. Sometimes the speakers relocate with their languages and settle permanently in a new location.
- viii. Violence resulting from different kinds of conflict, insurgency and other acts of terrorism in recent times in Nigeria is a fundamental candidate that is increasingly yielding to forced mobility of an entire community and sometimes an entire language group, the smaller groups in particular. Conflict as well as insurgency in Nigeria is characterised by burning of homesteads and other properties, mass killing, reprisal attacks and other acts that are meant to evoke terror on the victims. The result is often the forced movement of people from their native home towns to other places.

Mobility of people, which inadvertently include mobility of languages, is not new to Nigerian history. Mobility of Hausa speakers in pre-independent Nigeria contributed significantly to the spread and dominance of Hausa language across the northern region (Last 1993). The impact of mobility can also be seen in the complex linguistic geography especially in the linguistically diverse North Central and North East Nigeria. For example, some language groups in North East Nigeria such as Jukun can be found in different places in Taraba, Benue and Plateau states. But the current increasing rate of forced mobility especially in North Central and North East Nigeria is yielding to a change in the linguistic ecology of the affected places. Sometimes the displacement involves the movement of an entire community of speakers to, in some cases, unidentified locations. The languages on the move are usually the minor languages with relatively smaller number of speakers – a few tens of thousands to hundreds of thousands. The displaced speakers usually stay for an undefined length of time in the temporary camps. The displaced speakers as well as their languages always find a temporary home among stronger and more dominant languages. This situation inevitably yields to language dominance and threat of extinction of the smaller languages on the move. It also means that the displaced people who are unable to speak the language(s) of the host will be unable to accomplish a fundamental need for every living thing, which is communication. The dimension of forced mobility of language groups owing to violent conflicts and terrorism is yielding a situation that has significant impact on all aspects of development in Nigeria and some neighbouring countries such as Cameroon but more specifically on the cultural heritage and identity of a cross section of people.

4. Aims, methods and scope of the study

Crises – violent clashes, insurgency, militancy, etc. – often recur in Nigeria owing to inability to garner adequate information to make for holistic approach to the situation. One of the outcomes of crises is movement of victims and other people away from the crises area. In scholarship and in secular domains, the issue of violence has received significant attention. However, written correspondences, conversations, media reports – in digital and physical spaces – do not often take cognisance of the effect of high rate of forced mobility resulting from violence on the linguistic ecology of Nigeria. The data for the study reflect the dilemma of a people who are cut off from communication and information grid. It also highlights the precarious situation of the languages on the move, which embodies the worldview of the people who are floating. Information from this study will be of relevance to the society, to scholarship on African linguistics and to all humans who understand what language as well as communication means to humanity.

The attention of this work is on the impact of forced mobility on people, their language(s) and on Nigeria's luxuriant linguistic ecology. The study specifically examines the linguistic repertoires of a group of refugees designated in Nigeria's parlance as Internally Displaced Persons (IDP), who have been forced to move from their ancestral homes owing to violence as well as other acts of terrorism into a new community (IDP camps) since 2014.

In conversations relating to the impact of terrorism and violent conflicts on national development, the different arms of government in Nigeria conspicuously exclude or ignore the language component. Different humanitarian groups with varied interest such as social, political, religious, medical, etc., visit the camps to provide sundry aids to the IDPs. According to the IDPs, no group before this study has focused attention on their languages. Their basic language needs, which will ensure their integration into the immediate host community and to the larger society, are often not given adequate attention. Scholarship on IDP is often focused on economic and social issues, which do not include linguistic variables and the effect of the displacement of languages on Nigeria's linguistic heritage. The attempt of this study to investigate the linguistic repertoires of the IDPs in the camps becomes crucial. Information on how the languages on the move as well as their speakers are faring in the new locations and the resultant sociolinguistic trends serve to fill the identified gap in scholarship.

The study area is North East Nigeria, a space that has been characterized by violence orchestrated by the activities of insurgents, which has led to the displacement of several persons and communities in the area (Refworld 2015). Thus, there are several IDP camps in North East states – Borno, Yobe, Gombe, Adamawa, Bauchi and Taraba – and in some North Central states in Nigeria. The different language groups in the six states in North East Nigeria have been involved, in recent times, in different kinds of violent conflicts and terrorist attacks – Boko Haram insurgency, communal clashes, farmers/herdsmen clashes, and religious crises, among others. The persons in some of the camps were displaced owing to farmer herdsman clashes but the focus of this study is on IDPs who have been displaced owing to the activities of the insurgents in North East Nigeria. Camps for the study were selected using the stratified random sampling technique.

The study adopted a sociolinguistic orientation to reflect the current developments in the ecological landscape of languages in Nigeria. It specifically used some principles

from Ecolinguistics to reflect forced mobility as a salient factor that bears on the rich linguistic ecology of North East Nigeria. It used narratives from fieldwork, which was carried out in two selected IDP camps in Adamawa state in June 2018, to highlight the connection between forced mobility, language starvation, language endangerment and language maintenance. It basically reflected the experiences of persons who have been forcefully ejected from their homes and labelled as IDPs. Some of the displaced persons now find themselves in new environments, where they suffer physical starvation as well as language starvation as a result of inability to express themselves. The study employed the qualitative research method. The primary instrument of data collection in qualitative research include: observation, interview, focus group discussion, artefacts (records, documents, etc.), specimen records, etc. The data for this study were derived through observation, oral interviews and focus group discussions. The essence of employing different instrument for data collection is in the attempt to capture the nuances of the expressions and emotions of the informants.

A general survey of the language use spectrum in the camp was conducted in June 2018. The linguistic practices of the people in various domains were observed and recorded with both audio and video devices. Additional information and explanations were derived through personal interviews and group discussions. The interviews were conducted in English with Hausa interpreters. Wengraf (2001) model of interview which involves the use of elicitation and provocation of story was adopted to derive narratives of first-hand experiences from the subjects. Some of the people who served as informants in the interview sessions were the camp officials, women group leaders, young men and teenage girls.

In addition, there were focus group discussions with women, teenagers and children who could communicate intelligibly in Hausa or English. The discussions served to provide additional information and clarifications on the linguistic repertoires of the people in the camps. In Saint Theresa IDP camp Jimeta Yola, some women, about forty in number, who are direct victims of Boko Haram attack in Mubi and Michika areas also provided information in a group discussion. The women were invited by Saint Theresa Chaplaincy on June 19, 2018 to receive aids from donors. The women are all widows. Their husbands and older sons were massacred in the attack. With the restoration of relative peace in their area, they have moved back to their communities.

5. Conflict and mobility of languages in North-East Nigeria

Violent clashes especially in the northern part of Nigeria are, in fact, part of the history of hundreds of language groups and communities who inhabit the geographical space. Several communities in north-east region have history of ethnic clashes over issues ranging from land disputes to politics and power tussles (Mgbemena 2015). However, in Nigeria's recent history, the North-East is characterised by insurgency from a religious group, Boko Haram, which has resulted in the death of hundreds of thousands as well as the destruction of infrastructure in the whole region (*Amnesty International Report* 2015).

Generally violent clashes in the zone – ethno-religious conflicts, inter community clashes, insurgency – are marked by looting and massive destruction of lives and properties. Destruction of properties include the burning of farmlands and houses. Thus, during such

crises, people in the affected communities involved – especially women and children – often flee from their homes to settle temporarily or permanently in other areas, depending on the nature and duration of the clash. This factor has made significant and profound impact on the flora and fauna as well as linguistic topography of Nigeria. According to *The Punch* (2019) news report on a recent attack in Nigeria, about 1.8 million people are still homeless.

The dimension of violence in the case of insurgency and farmer herdsmen clashes appears deeper as it reportedly involves the use of sophisticated weapons of mass destruction, kidnapping of women and teenage girls, gruesome hacking to death of several individuals including children and pregnant women to the extent that the survivors who witness such onslaught of terror often flee from their ancestral homes in different directions. Sometimes the journey for survival takes them outside the boundaries of Nigeria especially to Cameroon (*Amnesty International Report* 2015).

The study population comprises displaced persons who have witnessed brute terror in their communities and are survivors of the horror. Some lost family members and friends, some of the elderly women in the group witnessed the killings of their husbands and their sons as well as the abduction of their daughters. Among the displaced persons one can find a lone survivor in the entire family.

5.1. Language and conflict in Borno state

Borno is the farthest state in north-east Nigeria. It shares boundaries with three countries – Niger, Cameroun and the Republic of Chad. The state, which is said to have housed over thirty six clearly identified language groups before the insurgency, has experienced severe attacks from Boko Haram insurgency so far. The state is close to the infamous Zambisa forest, the headquarters of Boko Haram, a group which gained the attention of global media with the kidnapping of about 300 school girls in Chibok, a community in the state.

The activities of Boko Haram have been heavy in Borno with the capital in Maiduguri. The parts of the state that have been seriously affected by the activities of the group include linguistically diverse communities in Bama and Gwoza, which are located close to the Zambisa forest and share boundaries with Adamawa state and Cameroon.

Gwoza came under heavy attack by Boko Haram in 2014 (*Leadership* 2014). The bouts of attack led to the massacre of several hundreds of people – predominantly the males, including the traditional ruler of the place, the Emir of Gwoza (see *This Day* 2014). The survivors fled to Cameroon and some to Adamawa state. The attacks involved massive killings with deadly weapons such as bombs, burning and destruction of properties in the communities and the establishment of some parts of the communities as Boko Harm territories. Although the Nigerian military has reclaimed the communities, the people still consider the places vulnerable to further attacks, hence their continual stay at IDP camps while their communities remain desolate. The population of Gwoza before the attacks is said to be about 300, 000 people.

6. Sociolinguistic situation in the IDP camps

The IDP camps in Adamawa state were established to accommodate victims of Boko Harams' multiple attacks in some parts of Adamawa and some communities in Gwoza (Borno state) in 2014 (*Reliefweb* 2014 report). The attacks led to massive movement of people into Yola. Some of the camps are still functional and inhabited by a large number of displaced people, some have only very few people still living there, while some have been shut down. The study selected camps that are still functional with relatively large population of IDPs from linguistically diverse communities in Adamawa and Borno states – Saint Theresa Catholic Church Internally Displaced Peoples camp, Jimeta Yola and Internally Displaced Peoples camp Damare, Girei Local Government.

6.1. Linguistic background of the IDPs in the camps

The population under study in the camps comprises people from East Gwoza and few people from communities in Madagali, Gulak, Basa and Michika and areas around Mubi local government area in Adamawa state. With the restoration of relative peace in Adamawa state, a significant percentage of the displaced persons returned to their homes, while displaced persons from Gwoza area await the restoration of lasting security in their communities.

A source who provided information at St. Theresa IDP camp in Jimeta Yola narrated that some of the people who fled the areas under attack arrived in Yola and were roaming the streets. In response Saint Theresa Catholic church provided accommodation for the people who took refuge in the church premises irrespective of their religion and denomination. With time, parishioners, individuals and organisations began to send aids. Years after the attack, the church continues to provide additional accommodation for the camp, which is still located within the church premises while the agencies such as United Nations Refugee Agency – UNHCR, Doctors Without Borders, as well as individual and corporate bodies provide sundry supports – social, religious, financial, medical, etc. The IDP camp in Damare, Girei, Adamawa state was established by Adamawa State Emergency Management Agency in response to the influx of people roaming Yola town at the heat of the crises in 2014 at the National Youth Service Orientation Camp. With the support of UNHCR, the people were able to raise tents for accommodation around the area. Government, non-governmental and international agencies as well as individual provide sundry support in the camps.

The languages of the people of East Gwoza, Borno before the displacement include; Gava, Glavda, Chinene, Lipdi, Dagwada, Chikide, Tandara, Zalivda, Wandala/Mandara, Maffa, Ndaba, Waha, Guduf, Waga among others. Most of the people who provided information during the interview in the camps stated that they belong to the following language groups – Chinene, Lipdi, Gava, Glavda, Zalivda. Chikide, Dagwada. Interaction with some camp officials reveals that other languages found in the camps include Hausa, English, Fulfulde, Michika, Fila and Bura.

6.2. Language use spectrum and the daily linguistic repertoires of the persons

As stated above, the IDPs in the camp are from diverse linguistic backgrounds – the diverse language groups that were displaced moved with their indigenous languages to the camps. This entails the existence and daily use of several languages in each camp. Out of these several languages, two – Hausa and English – serve as the general languages of communication.

The IDPs generally have histories of multilingualism practice as a significant number, especially the males, can speak more than one language. It is common to find a child, especially of school age, who uses more than three languages at once. But, a significant number of the displaced people from East Gwoza, especially the older women, do not speak Hausa, which serves as the *lingua franca* in North East Nigeria and the national language, English. Women in this category can speak one or two other languages in contact with their mother tongue, which sometimes share some features with the languages in contact. Because such women do not understand the general languages of communication in the camps, they constantly depend on others who can understand their language to receive information and to communicate their concerns in the camps.

Hausa plays prominent functions in their new home – the IDP camps. It serves as the language of general assembly as well as the language of address, announcements, inter group communication and instructions on various skill acquisitions and trainings. Hausa performs a lot of functions and has acquired the status of a dominant language within and outside the camps, in the general society where it serves as the *lingua franca*.

The English language, which is the national language in Nigeria, is used in the IDP camps as the language of documentation. All formal documentations and records are kept in English. It is used by bodies, agencies (national and international) who visit the camps for sundry reasons, which include rendering medical supports and other forms of aids. Within and outside the camp, English is the major language of health care and medical services. Even when the workers in the medical field use Hausa, a large volume of the register of medicine consists of the English language vocabulary so that the ability to use the language is evidently of an advantage in the community.

Hausa plays major role in religious activities in the camps. Hausa and English serve as the language of Christian worship. The language of homily for the Catholics include Hausa, English and Latin. English and Hausa choruses are used during worship. During prayers, some of the people leading prayers in Hausa often intersperse prayers in Hausa with English expressions. English and Hausa (the various varieties of English and Nigerian Pidgin) also serve as the language of announcement, communication and information dissemination. Hausa and Arabic are used by the Muslims in the mosque. The English language plays no role in Islamic worship and prayers.

The English language and Hausa language function as the general language of communication and information dissemination in the IDP camps. But English has a higher status than Hausa as it is the national language in Nigeria. It is the language of formal education. School children learn English in schools. A significant number of the displaced people in the camps, excluding school children claimed that they do not possess adequate facility in English.

Other diverse indigenous languages in the camp – Chinine, Glavda, Lipdi– are basically used as a medium in group communication, interaction and identification by the speakers. Some of the languages have very few speakers in the camp. Some of the speakers could be in various other camps within and outside the country. Some responses from the subjects indicate the importance they attach to their languages. Equally, some of the displaced persons shared their concerns about the survival of their language and people generally.

A young teenage girl narrated in an oral interview in the camp at Damare, Girei in June 18, 2018,

... Camp life is depressing but what gives me relief is when I interact with my friends here who speak my language, Glavda, it makes me feel that I belong somewhere and soothes the feeling that I often have of being lost. I often have the feeling that I am lost. My greatest fear is that I may not be able to go home again. Speaking my language gives me hope.

A camp official in Damare camp explained that the major challenge in the camp is that the displaced people usually create groups for themselves along ethnic lines. He claimed that religion is not considered a barrier in the camp as the people have common experiences that have bridged religious lines. But language and identity are important factors as the IDPs cling to other people who share the same language with them even when it compromises the unity of the camp.

6.3. Sociolinguistic issues from the linguistic repertoires of the IDPs

From the linguistic practices of the displaced people in the camp, it is evident that the speakers of the minor language in the camp would have some challenges. Those who cannot use English and Hausa encounter some challenges in the areas of self-expression, financial independence, among others. Basically, the scenario above presents some peculiar sociolinguistics issues which are presented below.

Language politics is one of the outcomes of the linguistic practices of the IDPs. Nigeria is a linguistically diverse nation with myriad languages of different status with no definite language policy. This situation will create various forms of language politics (Egbokhare 2004). Language politics has a significant impact on the linguistic topography of Nigeria. Language politics in this instance manifests in the struggle for dominance, relevance and survival of language(s) orchestrated either by the actions of the state or by the activities of the language groups.

The dominant languages – Hausa and English – continue to gain advantage over the minor languages used by the displaced people. Those who do not understand Hausa and English are disadvantaged – they basically depend on those who understand their language as well as Hausa or English. Sometimes, they are cut off from vital information and services.

In the course of the research, there was an incidence where some volunteer doctors visited Saint Theresa Catholic Church IDP Camp Jimeta Yola to offer free medical services. The doctors presented themselves as non-Hausa speakers. The management of the camp provided an interpreter, who relates information to the displaced persons in Hausa. But the interpreter could not serve those who do not speak Hausa. They were advised to wait, perhaps indefinitely, for someone who could understand their mother tongues and interpret their concerns at least in Hausa. If the person eventually did not return to meet

the medical team, they automatically forfeited the services. Inability to speak the dominant languages – Hausa and English – denies a displaced person fundamental human needs, the need to express oneself as well as the need to use language to access other facilities that can make for the wellbeing of any human being.

Apart from the need to express oneself, other socio-economic as well as cultural variables are integrated into linguistic practices in the camps. Ability to speak Hausa and English ensures faster integration into the host community. Those who can speak either of the languages or both easily move in and out of the camps. They can get jobs and make some money in the community. But those who cannot are limited and remain dependent on charity, which is often meagre. They are sometimes cut off from vital information and services that will facilitate their well-being and progress.

Ability to speak – Hausa and English – also attracts some measure of prestige. It attracts important assignments which sometimes have attached financial remuneration and higher social status. An informant explained that persons who cannot speak the languages are always isolated while those who can speak Hausa and can attempt a form of English are able to socialise. The displaced persons who can speak English assume some form of importance while the persons that are not able to use Hausa and English suffer some form of identity issues as they appear to be silenced and lost. While there is an attempt by the younger ones to use Hausa, the older women still cling to their language and identity. They remain isolated in the camps.

In a group discussion in Damare camp, the issue of communication challenges of the people who do not use Hausa and English was raised. In response, a participant explained that the management of the camp organises English language proficiency classes only for the adults as the younger ones are enrolled in formal education where they learn English and Hausa. A staff of an agency that provides counselling services added that the classes appear not to yield much result. Some of the women who speak Hausa and are in the English language class complained that the classes are not regular and that the contact periods are few. One of the instructors complained that the women are too disinterested to learn. A camp official added that the women have been so traumatized that nothing seems to arouse their interest anymore.

Language and identity shift within and outside the camp were part of the observations of some informants, in particular the women from Adamawa who are predominantly Michika speakers. They claimed that after the Boko Haram attack on their communities, which made them flee appeared to have abated, they returned to resettle. The women asserted that they are experiencing some form of challenges relating to language politics. According to their account, apart from Hausa, Fulfulde is considered important in their community and in some parts of Adamawa state where people in important positions are speakers of the language. The formally displaced women who are trying to reintegrate into the community explained that a good number of them are learning to use Fulfulde as it guarantees access to some provisions and aids from authorities.

Language endangerment is a significant outcome of the linguistic practices of the people. In all narratives presented in this study, the languages on the move as well as their displaced speakers are in a disadvantaged condition. The languages in East Gwoza have lost a significant percentage of the speakers with the few remaining still scattered in different places, with no concrete assurance of going back to their base and with no

prospect of thriving in the midst of dominant languages in contact. Some of the languages on the move are not being used. This poses a threat to their survival as the security situation in their communities is yet to stabilise (*Reliefweb* 2018).

7. Conclusion

Violent conflicts and acts of terrorism contribute significantly to forced mobility of persons and languages in Nigeria. Forced mobility has profound consequences not just on language but on the speakers who are grappling with a fundamental human need which borders on their inability to achieve communication.

The survivors of the attacks who are found in the camps are mostly women and children. Some of the women who witnessed the brutal killing of their husbands and sons have been consequently traumatized into silence. Some barely manage to speak even when they are able to understand and use the general languages of communication in the camps. This situation is made worse by the prevailing culture in north Nigeria that regards silence as a virtue and encourages women to speak sparingly. On the other hand, a good number of the children who also passed through the traumatic experience have been exposed to other languages in schools and in the host community generally. Some of them for the first time are being exposed to urban life as they were living in farmsteads in their communities. Some of the school children said that they would not want to go back to their communities again. To an extent, the children of the survivors are the ones that would ensure the continuity of the languages. The continuous stay of the children of speakers of the displaced languages in communities where the indigenous languages are rarely used and where they are exposed to other languages clearly poses a threat to the sustenance of the displaced languages. A good number of the internally displaced people in the camps under study have been away from languages since 2014. There is no concrete assurance of resettling back in the ancestral homes as fresh attacks are regularly launched in Borno (*The Punch* 2019). The indices clearly reflect that the languages on the move are endangered. The continuous displacement of languages as a result of violence would invariably yield a serious impact on Nigeria's rich linguistic ecology.

References

- Amnesty International Report* 2015. "Our job is to shoot, slaughter and kill: Boko Haram's reign of terror in Nigeria". www.amnesty.org. Retrieved March 12, 2018.
- Ballard, John. 1971. "Historical inferences from the linguistic geography of the Nigerian Middle Belt". *Africa* 41, 295-305.
- Blench, Roger. 2012. *An atlas of Nigerian languages*. Cambridge: Kay Williamson Educational Foundation.
- Crozier, David & Roger Blench. 1992. *An index of Nigerian Languages*. Texas and Abuja: SIL and the NERDC.
- Egbokhare, Francis. 2004. "Language and politics in Nigeria". In: Owolabi, Kola & Ademola Dasyilva (eds.). *Forms and function of English and indigenous languages in Nigeria: a festschrift in honour of Ayo Banjo*. Ibadan: Group Publishers, 507-522.
- Elugbe, Ben. 2009a. *Multilingualism, national languages, and lingua franca in Nigeria*. Ibadan: University of Ibadan.

- Elugbe, Ben. 2009b. Art, culture, language and national integration. A paper presented at the workshop of National Institute for Policy and Strategic Studies (NIPSS) Kuru.
- Ethnologue* 2018. *Languages of the world 12th edition*, www.ethnologue.com/country/NG. Retrieved on 1 September 2018.
- Harmon, David. 1996. "Losing species, Losing languages: Connection between biological and linguistic diversity". *Southwest Journal of Linguistics* 18, 89-109.
- Harnischfeger, Johannes, Rudolf Leger & Anne Storch. 2014. "Lower rank greets first: getting along in multilingual communities". In: Anna Storch, Johannes Harnischfeger & Rudolf Leger (eds.). *Fading delimitations multilingual settlements in a convergence area case studies from Nigeria*. Köln: Rüdiger Köppe Verlag, 1-36.
- Hatoss, Aniko. 2013. *Displacement, language maintenance and identity. Sudanese refugees in Australia*. Netherlands: John Benjamins.
- Last, Murray. 1993. "History as religion: de-constructing the Magians 'Maguzawa' of Nigeria Hausaland". In: Jean-Pierre Chretien (ed.). *L'invention religieuse en Afrique: histoire et religion en Afrique noire*. Paris: Karthala, 267-296.
- Leadership* 2014. "Nigeria: Gwoza Under Siege – Boko Haram Kills 300, Wipes Out Three Villages". <https://allafrica.com/stories/201406050381.html>. Retrieved on March 10, 2019.
- Meek, Charles Kingsley. 1931. *A Sudanese Kingdom: An ethnographic study of the Jukun speaking peoples of Nigeria*. London: Kegan Paul Trubner & Co.
- Mgbemena, Judith Adaku. 2015. "Perspectives on Multilingualism in Nigeria: a particular reference to some language groups in North East Nigeria". A paper presented at a conference on African Multilingualism – Motivation, Modality, Movement and Meaning, London 2015.
- Muhlhausler, Peter. 2003. *Language of environment, environment of language: a course in Ecolinguistics*. London: Battlebride.
- Nassenstein, Nico & Andrea Hollington (eds.). 2015. *Youth Language Practices in Africa and Beyond*. Berlin: Mouton de Gruyter.
- Nettle, Daniel & Suzanne Romaine. 2000. *Vanishing voices: the extinction of the World's languages*. New York: Oxford University Press.
- Refworld* 2015. *United Nations Refugee Agency (UNHCR) report. Country reports on terrorism 2014 – Nigeria*. <https://www.refworld.org/docid/5587c74628.html>. Retrieved on March 10, 2019.
- Reliefweb* 2014. *Nigeria: Update on IDPs in Camp and Host Communities in Adamawa State, Situation Report No. 1 (2 October 2014)*. <https://reliefweb.int/report/nigeria>. Retrieved on March 10, 2014.
- Reliefweb* 2018. *Médecins Sans Frontières. reliefweb.int/report/nigeria/situation-northeast-nigeria-deteriorating-after-years-conflict*. Retrieved March 10, 2019.
- Shimizu, Kiyoshi. 1980. *A Jukun Grammar. Beiträge zur Afrikanistik 9*: Institute für Afrikanistik und Ägyptologie der Universität Wien.
- Spencer, John (ed.). 1971. *The English language in West Africa*. London: Longman.
- Stibbe, Arran. 2015. *Ecolinguistics: Language, ecology and the stories we live by*. London: Routledge.
- Terralingua* 2008. *Terralingua: unity in biocultural diversity* (homepage). <http://www.terralingua.org>. Retrieved on January 18, 2019.
- The Punch* 2019. "Thousands flee Boko Haram attack in Borno". <https://punchng.com/thousands-flee-boko-haram-attack-in-borno/>. Retrieved on February 22, 2019.
- This Day* 2014. "Boko Haram Kills Emir of Gwoza in Highway Ambush". <https://allafrica.com/stories/201406020092.html>. Retrieved on March 10, 2019.
- Webster, James Bertin. 1993. *Kwararfa: The traditional face of the coin*. Boston: Boston University Centre for African Studies.
- Wengraf, Tom. 2001. *Qualitative research interviewing: Biographic narratives and semi-structured methods*. London: Sage.

The function of nonverbal regulators in Hausa face-to-face interaction

Abstract

The structure of conversation is a subject of many empirical researches along the Conversational Analytical point of view. Scholars have agreed that nonverbal signals which are called regulators act as a kind of traffic signals that aid in the flow of interactions. This paper studies the use of nonverbal regulators in Hausa social interactions, through the use of data consisting of natural face-to-face interactions taken from multiple settings. The research discovers that Hausa peoples' interaction, especially in face-to-face settings, is marred with different gestures and postures acting as conversational regulators. Beyond regulating the interaction, these gestures and postures contribute to meaning making and influence communication outcomes.

Keywords: Hausa, Nigeria, gestures, face-to-face interaction, regulators

1. Nonverbal regulators

In our everyday social interactions we communicate through verbal and nonverbal mode, which may complement, substitute, accentuate, repeat, regulate and stand in opposition to the verbal form of communication. Duncan (1972) maintains that conversation should be an activity organized in such a way that participants interact without bumping into each other. The organized way of communication requires the use of behavioral signals that serve as turn-taking mechanisms. Gestures or body postures that help us control and understand the conversation better, i.e. they regulate conversational flow are referred to as regulators (Ekman & Friesen 1969). Regulators include nodding, leaning forward and raising eyebrows among others. Regulators are sometimes referred to as turn-taking signals because they are used to decide who and when should speak. For example, we use hand signals to indicate that we finished talking; we may select the next speaker by using the gaze. Ekman (2004: 44) notes that there are signals the speaker makes to prevent others from speaking, such as 'holding a hand out' and he names them 'floor holders'. There are also 'turn seekers' responses that take place when the listener attempts to gain the floor. An example of a 'turn seeker' is the situation when a person almost raises from a chair.

2. Data and methodology

The data presented in this paper was gathered through video recording of natural face-to-face interactions that took place at school or during wedding ceremonies. Students of Aminu Kano College of Islamic Studies Kano were videotaped while interacting after their exams. The wedding ceremony took place in Kano Municipal in March 2017. The respondents were between the ages of 18 and 35. Most of the videos were taken by volunteers while the researcher remained a non-participant observer. Some respondents knew they were being recorded, others were taken unaware. However, even those aware of the recording did not know the purpose of the video. The length of the videos varies according to the duration of the interaction. The total time of videos is 30 minutes. In order to see how verbal and nonverbal forms intertwine in meaning making, the video recording was transcribed using Gail Jefferson's Transcription Notation provided in Mazeland (2006) and Jenks (2011). The data were scrutinized in accordance with Conversational Analysis (CA) (Sacks, Schegloff, et al. 1974). The main goal of CA is to describe the intertwined construction of practices, actions, activities and the overall structure of interaction. The symbols used for transcription are the following:

=	when one speaker follows the utterance of another speaker without a pause.
↑	rise in pitch
↓	fall in pitch
<u>word</u>	underlining signal salient stress
word:d	a colon renders a noticeable sound stretch
°	° softer utterance
(xxxxx)	unclear utterance
(.)	a short pause
(word)	description of nonverbal conduct is placed in parenthesis

3. Use of regulators

Regulation of interaction may be used by the speaker or by the hearer. The speaker uses hand gestures, vocalic cues and postures to explain, illustrate or emphasize what is being said verbally. The listener may in turn use nonverbal signals such as nodding, gaze, facial expressions and postural shifts as feedback mechanisms controlling the behavior of the speaker.

3.1. Regulators to mark the beginning of interaction

Nonverbal signals set the stage helping to start any face-to-face interaction or guiding the interlocutors on how to behave (Burgoon Guerrero & Floyd 2016). A Hausa proverb *shimfidār fuska ta fi shimfidār tabarma* 'a welcoming face is better than spreading a mat' indicate that from the facial expression of a host, the guest may understand whether he/she is welcome or not. Often, when we intend to interact with others, we do not form a verbal message like 'come on, let us sit and talk', but rather use nonverbal signals. Facial

expressions such as *kallo* ‘gaze’ or *murmushi* ‘smile’ as well as vocalic signals: *gyaran murya* ‘clearing the throat’ are conversational starters. However, their use depends on the context and interactional partner. For instance, same-sex friends encounter may just require the use of gaze to invite an interlocutor to start a conversation (Adamu 2019), but an elderly man willing to interact with a woman, especially a divorcee, may use *gyaran murya* ‘clearing the throat’ to make his presence known and to start a conversation (Adamu 2018). An example found in the collected data is using *murmushi* ‘smile’ as a conversational starter. The short interaction shown in Extract 1 was captured at the wedding event.



Fig 1. *murmushi* (smiling)

Extract (1). A short interaction between a woman and a girl. Background information: *A girl called Amira entered a room full of guests at the wedding ceremony; a woman (Anti) sitting inside the room smiled and the girl went straight to her.*

- 1 Anti: (gazes at Amira)
 2 Anti: a'a ke kadai↓? (while smiling)
 no 2SG.F alone
 ‘you are alone↓?’
 (.)
 3 Amira: (she kneels down, in front of Anti)
 4 Amira: eh Umma tana tahowa tana baya
 yes U. 3SG.F.IPFV coming 3SG.F.IPFV back
 ‘yes Umma is coming she is behind’
 (.)
 5 Amira: ina wuni?
 how spending.day
 ‘good afternoon’
 6 Anti: lafiya lau
 health IDEO
 ‘I am fine’
 7 Anti: ya makaranta ko ana hutu?
 how school Q 4.IPFV break
 ‘how is school or are you on holiday?’
 8 Anti: (she turns her head towards another person)
 9 Amira: (stands up and leaves the room)

Amira entered the room filled with many women at a wedding ceremonial sitting event, called *yini*¹. When she looked around she noticed the two signals used by Anti: gaze and smile that displayed the willingness to interact with Amira and in this respect served as a conversation opener. The use of these signals by Anti compelled Amira to go straight to her and kneel down to greet her. Although Anti is a friend of Amira's mother, the girl would not start talking to her without noticing these nonverbal signals.

3.2. Regulators to mark the end of interaction

There are signals that are known to be a part of conversational cessation ritual, which may be used with some degree of awareness by both the speaker and the listener. A speaker may decide to be silent or show a kind of readiness to go. The listener may use postural shifts, he may look at his watch or gaze somewhere. The example of marking the end of the conversation is shown in Extract 2.



Fig 2. *kawar-da-kai* (gazing away)

Extract (2). Conversation between two sisters. Background information: *Aisha waited patiently for her elder sister Bilki in school to get a book from her.*

- 10 Aisha: ga ni ina jiran ki
 COP 1SG 1SG.IPFV waiting-GEN 2SG.F
 'I am here waiting for you'
- 11 (Aisha walks towards Bilki)
- 12 Bilki: ai na faɗa miki sai gobe
 O.yes! 1SG.PF told 2SG.DAT till tomorrow
 'but I told you see you tomorrow'
 (.)
- 13 Aisha: (smiles)^oha:ba^oke kuwa[↑]
 but.no 2SG.F PART
 (smiles) '°but no^o please[↑]'
- 14 Bilki: (turns her head) ke kika sani.
 you 2SG.F.PFV.REL know
 'but you know'

¹*yini* is a wedding event that is carried out indoors. Normally, people sit on mats, chat and have fun while food is being served. This kind of setting is a female affair, so women interact freely with each other.

Aisha waited for Bilki to get a Hausa novel for her, but Bilki was not willing to give the book to Aisha. She told her to wait until the next day, but Aisha insisted on getting the book that day. Bilki performed the gesture called *kawar-da-kai* (turning the head aside), which may be interpreted in two ways: either she did not want to talk any longer or she was inconvenienced by Aisha's request to give her the book and she decided to terminate the interaction.

3.3. Regulators as turn-taking signals

Regulators help in making turn-taking providing an order within the interaction. By using gestures, postures and vocalic signals, interlocutors may show interest in taking over the floor, holding on to the talk stage, select the person to speak with or indicate readiness in giving up the floor. Hausa people use hand gestures to hold on to the talk or too draw the attention of the other interlocutor in order to emphasize a point. That is why hand gesturing is said to be largely, if not entirely, a speaker's phenomenon (Yang 2010). The excerpt 3 shows an example of a signal called *ware-yatsu* (finger spreading) used by Adama as a turn maintaining signal.



Fig 3. *ware-yatsu* (finger spreading)

Extract (3). An interaction between two students in front of the class. Background information: *Adama and Hauwa were revising their books just before the exam. They were discussing credit loads and the number of courses they were expected to take.*

- 15Adama: akwai elective a cikin gidan Hausa
COP elective PREP inside-GEN house-GEN Hausa
'there are elective courses in Hausa Department'
- 16 Hauwa: nawa ne a can? (pointing at her back)
how.many COP.M PREP there
'how many are there?'
- 17Adama: ashirin da biyar ne↑
twenty and five COP.M
'there are twenty five↑'
(.)
- 18 Hauwa: ai kosis din sha takwas ne
o.yes! courses DEF eighteen COP.M
'these courses are (for) eighteen'
(.)

- 19 Hauwa: a'a (raises her hand up and spreads her fingers while looking in front)
 'no'
- 20 Adama: ki dai lissafa↓
 2SG.F.SBJV PART count
 'why don't you count'

Adama mentions that there are courses worth 25 credits in Hausa Department and her interlocutor disagrees with her saying that they are worth 18, which is most unlikely. In order to emphasize her point Adama does *ware-yatsu* (finger spreading). The gesture also gives her a chance to hold the floor. Spreading the fingers performed by Adama may be interpreted as a beat. It is devoid of any semantic content, but may play a pragmatic function (Holler & Beattle 2002: 31). By spreading the fingers and holding the hand horizontally with the palm directed inward, Adama wants to stick to her claim that the courses have twenty five credit loads. Hauwa who notices that Adama is totally in control of the floor, waits for her turn. Meanwhile, Adama makes a signal – placing her hand on her lap to display her readiness to hand over the turn to Hauwa.

Extract (4). An interaction between four friends. Background information: *Four friends were interacting in school after writing an exam. Ummi who happened to be the one recording the video, and Rabi'a missed their exam and they were trying to put the blame on their friends Binta and Hadiza:*



Fig 4. *zumbura-baki* (pouting; inverting the mouth in an upside down “u” shape)

- 21 Rabi'a: Binta (.) ba ki da mutunci ko ki yi mana waya ba mu yi jarrabawa ba.
 B. NEG 2SG.F with humanity PART 2SG.F do 1PL.DAT telephone NEG
 1PL.F do exam NEG
 'Binta (.) you are irresponsible you did not call us we missed our exams'
- 22 Hadiza: (xxxxx)
- 23 Binta: (opens her mouth)
- 24 Binta: aka ce kun yi↑ (she widens her pupils) =
 4.PFV.REL say 2.PL.PFV do
 'they said you did it'
- 25 Hadiza: = aka ce kun yi
 4.PFV.REL. say 2.PL.PFV do
 'they said you did it'
- 26 Binta: (tilted her head)
- 27 Hadiza: Ummi aka ce a jira sai sha daya za mu shiga, kuma aka ce Rabia↑ ta/yi

- U. 4.PFV.REL say 4.SBJV wait till eleven FUT 1PL enter PART 4PFV.REL say R.
3SG.F.PFV do
'Ummi they said to wait till eleven to enter and I was told Rabi'a did it'
(.)
- 28 Binta: Rabi'a↑(she lowered her face, frowned and pointed to the place with her hand)
ni kuma a cikin nan na gan ki
R. 1SG 1SG PREP inside-GEN here 1SG.PFV see 2SG.F
'but I saw you there'
- 29 Rabi'a: ahhhhh
- 30 Binta: (opens her mouth)
- 31 Hadiza: (pouts and turns her head aside)

The interactional flow in Extract 4 is heavily based on gestures. When Rabi'a starts blaming Binta for not telling her about the exam, Hadiza whispers something. Binta who can hear it opens her mouth widely. Opening mouth in Hausa culture is an indication of surprise. In Extract 4 it is followed by a verbal comment uttered in chorus by Binta and Hadiza: 'they said you did it'. When uttering this fragment Binta opens her eyes very widely performing a gesture called *zare-ido*. *Zare-ido* usually expresses surprise or fear. In this instance, its meaning is not explicit. Perhaps she is surprised and vexed at the same time being conscious of the fact that her classmate missed an exam and now she is the one to be blamed. Binta tilted her head a bit, showing a sign of helplessness and willingness to give up her turn. Hadiza used the opportunity and took up the turn saying that we were told to wait till 11 to enter and they were also told that Rabi'a had written the exam'. Binta claims a turn and utters the name 'Rabi'a' with a high pitch and then she lowers her frowned face as an indication of helplessness. She also points at the place of exam saying: 'but I saw you there'. Rabi'a utters a vocalic signal 'ahhhhh' used when someone is in dire situation. As a result of the vocalic signal uttered by Rabi'a, Binta opens her mouth once again. Hadiza pouts and turns her head to the side (*kawar-da-kai*) performing the same gesture as shown in Figure 2. Pouting called *zumbura-baki* (the mouth is inverted in an upside down "u" shape) is used to replace speech and expresses condemnation, anger, disagreement, disregard or just pretending to be ill tempered. It is often accompanied by *kawar-da-kai* (turning the head aside). Hadiza used the gestures to indicate condemnation over what happened. She was also not happy about Rabi'a trying to put the blame on them, and she was conscious of the fact that the attempt to show their guiltlessness proved abortive.

Extract (5). Background information: *The event took place at school among the men who happen to be course mates. Anas, the class representative and Bala were searching for course material. Anas finally got hold of the material, made copies and distributed them to some of his class members. Bala was not given a copy.*

- 32 Anas: kana can kana bacci
2SG.IPFV there 2SG.IPFV sleeping
'you are sleeping'
- 33 Bala: (put his hand on his waist called *kama kugu* shifting his weight on one leg while listening attentively to Anas)



Fig 5. *kama kugu* (akimbo) and *kama-gemu* (beard stroking)

- 34 Hassan: me ya faru?
 what 3SG.M.PFV happen
 ‘what happened?’
- 35 Anas: (stroked his beard)
- 36 Anas: yanzu akwai wani za mu ba ka
 now COP certain FUT 1PL give 2SG.M
 ‘now there is one we will give you’
- 37 Anas: (removed his hand from his face and pointed to the right)
- 38 Bala: (released his waist and stood straight)
- 39 Bala: an yi baccin sai ka ce ba tare muka nemo ba
 4.PFV do sleeping until 2SG.M.PFV say NEG together 1PL.PFV.REL search
 NEG
 ‘I slept as if we were not looking [for it] together’

Bala came to school late only to discover that Anas got the material they had struggled to get, made copies and distributed them all without reserving one copy for him. Bala performed *kama-kugu* (akimbo) by placing hand on the hip. By doing one hand akimbo gesture and staring at the speaker, he indicates his attentiveness and not being ready to take a turn. He also shows that he has interest in the ongoing discussion. When Anas is asked by another man about what happened he performs *kama-gemu* (beard stroking gesture) followed by a short pause. *Kama-gemu* is purely a male gesture used by the listener as a feedback mechanism. The gesture is made by caressing one’s beard with the fingers. The beard stroking indicates that Anas is thinking hard and fast about the situation at hand. It can also be an attempt to yield his turn to Bala who is, however, unwilling to take it. The moment Bala is ready to take the turn, he takes a different posture – he stands straight releasing his hand from his waist; it serves as a turn-seeker response.

4. Conclusion

The paper shows the significance of nonverbal signals in Hausa social setting. Non-verbal signals continue to negotiate meaning and structure of verbal exchange. However, nonverbal regulators cannot be studied out of context. One has to consider many factors when deciphering such kind of signals: the relationship between the participants, previous interaction that may affect the meaning of the current interaction, physical setting and

finally the mood and temperament of the interlocutors. Depending on the communication context, regulators can be used as a substitute for verbal expression or as an emphasis of a verbal statement. However, one may choose to rely on verbal messages to serve the same purpose of interactional regulation, which could unequivocally replace gestural regulators in some given instances though hardly the vocalic ones.

References

- Adamu, Aisha Umar. 2018. Study of Nonverbal Communication in Hausa, PhD Thesis, Department of Linguistics, Bayero University, Kano.
- Adamu, Aisha Umar. 2019. "The Study of Oculistics in Hausa Nonverbal Communication". *Bayero Journal of Linguistics* 6(1), 24-33.
- Burgoon, K. Judee, Laura K. Guerrero & Kory Floyd. 2016. *Nonverbal Communication*. New York: Routledge.
- Duncan, Starcey 1972. "Some Signals and Rules for Taking Speaking Turns in Conversations". *Journal of Personality and Social Psychology* 23(2), 283-292.
- Ekman, Paul. 2004. "Emotional and Conversational Nonverbal Signals". In: Jesus M. Larrazabal & Lusi A. Prez Miranda (eds.). *Language, knowledge and representation: proceedings of the Sixth International Colloquium on Cognitive Science (ICCS-99)*. Dordrecht/London: Kluwer Academic Publishers, 39-50.
- Ekman, Paul & Wallace V. Friesen. 1969. "The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding". *Semiotica* 1(1), 49-98.
- Goffman, Erving. 1955. "On Face-work: An Analysis of Ritual Elements in Social Interaction". *Psychiatry* 18 (3), 213-231.
- Guerrero, Laura & Kory Floyd. 2006. *Nonverbal Communication in Close Relationships*. New Jersey: Lawrence Erlbaum Associates Inc. Publishers.
- Holler, Judith & Geoffrey Beattle. 2002. "A Micro-analytic Investigation of How Iconic Gestures and Speech Represent Core Semantic Features in Talks". *Semiotica* 142, 31-69.
- Jenks, Christopher. 2011. *Transcribing Talk and Interaction: Issue in the Representation of Communication Data*. Amsterdam: John Benjamin Publishing Company.
- Mazeland, Harrie. 2006. "Conversational Analysis". In: Keith Brown (ed.). *Encyclopedia of Language and Linguistics*, vol. 3. Oxford/Amsterdam: Elsevier, 153-162.
- Sacks, Harvey, Emmanuel A. Schegloff & Gail Jefferson. 1974. "A Simplest Systematics for the Organization of Turn-taking for Conversation". *Language* 50(4), 696-735.
- Yang, Ping. 2010. "Nonverbal Gender Differences: Examining Gestures of University-Educated Mandarin Chinese Speakers". *Text & Talk* 3(30), 333-357.

“Allah expressions” as a manifestation of common cultural area in West Africa

Abstract

Many languages of West Africa have adopted religious phrases from Arabic. They are used either in the form of direct borrowings (*assalamu alaikum*) or as their translated versions. The paper investigates the scale of penetrating the lexicon of some main West African languages by adopting religious phrases of Arabic origin and presents a more detailed analysis of their function in the Hausa language. The examples of the so-called Allah expressions contribute to the idea of common culture-linguistic area in West Africa. The available data from Hausa are to demonstrate the process of passing the sacred vocabulary and phraseology of Arabic origin on to the colloquial language and written literature.

Keywords: religious phrases, secularization of language, Hausa

1. Introduction

West African languages are saturated with a rich variety of expressions of Arabic origin which have mostly religious connotation. The words or multiword units are commonly used in the Arabic-speaking world, as for example:

- (1) a. *Assalamu alaikum*¹ السلام عليكم ‘Peace be upon you’.
- b. *Wallah* والله ‘[I promise] by Allah’
- c. *Alhamdulillah* الحمد لله ‘Praise to Allah’
- d. *Allahu akbar* الله أكبر ‘Allah is great (greater)’

They are identified as borrowings adopted by the speakers of African languages with the introduction of Islam. The iconic status they achieved in Arabic language was conducive to the increase of their communicative function and making them recognizable also among non-Arabic speakers. Most of these phrases are originally expressions invoking Allah, but similarly to Arabic, their use is not limited to religious discourse, they also function in ordinary conversation, irrespective of the religious affiliation (or otherwise) of the speaker. The so-called Allah expressions (Clift & Helani 2010: 358) are present in all forms of communication, both spoken and written.

¹ Transcription is based on ALA-LC convention for romanization.

The aim of the paper is to show how deeply the Arabic phraseology have penetrated the structural and communicative properties of West African languages. A more detailed analysis of the Hausa examples is to demonstrate the process of passing the sacred vocabulary and phraseology of Arabic origin on to the colloquial language and written literature.

1. Arabic borrowings as a culturally motivated innovation in African languages

The contacts of West African languages with Arabic, the language of Arabo-Islamic culture, have been maintained since at least the 12th/13th century (Löhr & Awagana 2007: 12). Arabic expressions invaded the African languages when Islam started to penetrate particular regions. For example, the country of Wolof have been influenced by Muslim religion and Islamic social organizations since the 16th century. In the 19th century, all its social institutions functioned on the basis of the Islamic law (Baldi 2003: 2). The Islami-zation of Sub-Saharan region of West Africa was conducted mainly through the powerful regional empires (Mali, Songhai, Kanem Bornu and Hausa) which were open to influences of the Muslim world and the Arabic language. The most dynamic *lingua franca* of a vast territory in this region was Hausa.

However, the cultural landscape of West Africa was shaped in mutual relations between African societies. The transmission of Islamic religion and cultural values through African cultures created a big area of convergence in West Africa, in which shared Arabic borrowings became differentiated according to cultural regionalisms.

There is a great number of words and phrases of Arabic origin in languages of West Africa. Following the most comprehensive source (Baldi 2008), 3034 Arabic items find their reflexes in various languages of West Africa and/or in Swahili, e.g. Ar. *sūg* ‘market’ is *sugu* in Bambara, *kasíwu* in Kanuri, *kàsuwā* in Hausa, *súugine* in Bidiya (Baldi 2008: 263). Their form does not indicate direct borrowings from Classical Arabic, which rather suggests the transmission through spoken form of communication. The most widespread loans occur in languages such as Wolof, Kanuri, Songhay, Hausa (Baldi 2004: 85). This confirms that the cultural vocabulary (including religious terms and phraseology) was acquired through local institutions and main languages of the area which had the function of *lingua franca*.

In this context, Hausa is of special significance as a vehicle in the distribution of Arabic lexicon and phraseology among the languages of West Africa. A large number of Arabic borrowings in Hausa is confirmed in many works (Greenberg 1947; Yalwa 1992; Baldi 2008). It reflects the communicative role of this language in the region and makes it the basis for studies on their further development.²

Small languages were also heavily influenced by Arabic. Multilingualism enabled passing the loans from one language to another, but the prestige of Arabic, which was “favoured by its charm and by its status symbol” (Baldi 2004: 72), was also important in absorption

² The political power of the Hausa states was an important factor of the language (Salamone 2010), but this status was also determined by the fact that the Hausa people were commonly recognized in West Africa as good traders and zealous Muslims, therefore their language was and have been easily absorbed into other languages.

of new terms into the indigenous languages. The contact of the members of West African population with spoken Arabic was also facilitated by their travels to Arabic-speaking countries, mostly to Sudan.

Referring to the main source of the present data, which is based on the existing dictionaries (Baldi 2008), religious phrases invoking Allah belong to the most widely dispersed forms of Arabic expressions in African languages (*wa-llāh*, *al-ḥamdu lillāh*, *as-salāmu 'alaikum* has been noted for all main languages of West Africa, including Wolof, Songhay, Fulfulde, Hausa, Bambara).

2. Incorporation of Arabic phrases into West African languages

The Arabic phrases which function in African languages have different variants. Their written form follows either the pronunciation (including marking tones) or the standard orthography rules in Latin script. Their meaning and use may be somehow different from the original Arabic forms, but they have much in common in respect of their communicative functions and are recognizable in the whole area as a common cultural code. Here are some examples of phonological variants of the Arabic phrases in West African languages and the most salient characteristics of their use in non-religious discourse.

The phrase *inshallah* [in šā'a allāh] *إن شاء الله* 'God willing' (lit. if Allah wills) also in Arabic has many variants and cognate forms, such as *inshaalla*, *inshalla*, *insha'Allah*; it occurs in a variety of forms, sequential combinations and has many communicative functions. For example, when used at the beginning of the utterance, *inshallah* marks "the expression of hope for a desired outcome", whereas its postpositional use is to denote the termination of one topic and moving to another (Clift & Helani 2010: 363f).

In languages of West Africa, the following forms have been identified (Baldi 2008: 285):

- (2) a. Hausa: *in shā àllāhù*
 b. Kanuri: *ínshállâw*
 c. Songhai: *ìnsàllà*
 inšaalla, inšaallah
 d. Wolof: *insàllaa, sàllaaw, sàllaaw*

African languages confirm the use of *inshallah* in contexts referring to future actions. However, in contextual use, it is rather the meaning of confirmation, agreement which is more clearly manifested. In Songhai, confirming the statement ('without any doubt') is given as its equivalent (Baldi 2008: 285). In the Kanuri version of modern literary text *Amina* the phrase *Insha Allah* is used to confirm the farewell statement, e.g. in the dialog (Umarbe 2014: 1):

- (3) A: *Nyiga tusshiya ruki* 'I'll see you later'
 B: *Insha Allah.* (Ok, fine, I'll be looking forward to seeing you)³

³ In the English version of the novel (Umar 2005) the response to 'I'll see you later' is given as *In sha Allah*.

The sequential context, in which the expression occurs, varies. *Inshallah* may be used at the beginning of the utterance, but it may also follow it. In Hausa, in postpositional use it has gained the status of a phrase modifier, as in *matarka in sha Allahu* ‘your wife-to-be’.

Another phrase *wa-llāh* ‘By Allah’ (Arabic *والله*) is used to express a great credibility on an expression. It also functions as a promise to do something. Often translated as ‘swear to Allah’ it also means ‘I promise by the one who holds my soul, i.e. Allah’. In African languages, the reference to the form *wallaahi*, i.e. ‘by my Allah’ (often contracted to *wal-lai*) is more common, e.g. (Baldi 2008: 497):

- (4) a. Hausa: *wallāhi! wallà!*
- b. Bambara: *wàlayi!*
- c. Songhay: *wallaahi! allaahi!*
- d. Wolof: *wallaay*
- e. Dangaleat: *wállày*

In Hausa, *wallahi* is often combined with the other interjections having the form of Allah expressions or the word Allah (*Alla*) itself, e.g. *wallahi, Alla, gaskiya ne* ‘it is really true’ (Abraham 1962: 919). *Wallahi tallahi!* is basically an exclamation (‘Great God!’) but it is often added to other statements to make a promise or to confirm the real existence of what is being said. In this function, as an equivalent of ‘really’ it makes the conversation more emotional.

The formula *as-salāmu ‘alaikum* *السلام عليكم* ‘peace be upon you’ is widely used in Arabic as a common way of greeting. It functions as an equivalent of ‘hi, hello, good morning/ afternoon/ evening’. It can be also used in farewells. The phrase is modified, depending on some additional elements and whether the formula is addressed to one or several persons. The variants also function in the phrases used in Africa, as in the following examples (Baldi 2008: 252):

- (5) a. Hausa: *sàlāmà àlaikùn!* (*sàlāmà àlaikà* if one person is greeted)
- b. Kanuri: *salámalékum!*
- c. Songhai: *salaamuwaleykum!*
wassalaamaleykum!
- d. Wolof: *salam alekum!; wasalaam!*
asalaamalaikum⁴

As a reply, the phrase *wa ‘alaykum as-salām* *و عليكم السلام* (‘peace be upon you’) is used in Arabic. In African languages, it may be also modified and different forms function, e.g.:

- (6) a. Hausa: *àlaikùn sàlāmù* (if more persons than one are greeted)
àlaikà sàlāmù (when said to one person)⁵
- b. Maninka: *àlekunsalam*
- c. Songhai: *aleykummsalam!*

⁴ Following Bichler (2000: 231).

⁵ A variant form is *àssalāmù!*

- aleykumwassalam!*
waleykummasalam!
- d. Wolof: *malaikum salaam*⁶

Along with a variety of forms, the social context for the appropriate use of this greeting has changed recently. As a greeting formula of the Muslim people for a long time it has been restricted to men.⁷ Nowadays, the formula is also widely used by women in the whole Sahel area. It is also accepted as a form of greeting used by the members of non-Muslim societies and foreigners.

The Arabic expression *al-ḥamdu lillāh* الحمد لله ‘Praise/thanks be to Allah’, is usually used to express satisfaction, or after having finished eating, or if someone asks you how you were and you want to say that you are satisfied.

In African languages, it is added mostly as an interjection or an additional statement to express satisfaction (‘thank Allah’). Its form and spelling vary, e.g. (Baldi 2008: 152):

- (7) a. Fulfulde *alhamdu lillaahi*
 b. Hausa *àlhamdù lillāhī*⁸
 c. Kanuri *alhámdu lillahi*
 alhámdu llai
 d. Maninka *álahamudulillay*
 alíhamudilillahi
 álihamdulillayi
 e. Songhay *alhamdilillaahi*
 alhamdu

The sample dialog in Wolof (Bichler 2000: 231) confirms the use of the phrase *alhamdulillahi* in conventional greetings, e.g.:

- (8) – *Ana xale yi?* ‘How are your children?’
 – *Jamm rékk, alhamdulillahi.* ‘They are fine, *alhamdulillahi.*’
 – *Mba kënn feebarul?* ‘I hope that you are well’ (lit. that nobody is sick)
 – *Deedeet, jamm rékk, alhamdulillahi.* ‘All are fine, *alhamdulillahi.*’

Phonological variants of the Arabic phrases in African languages are rather close to each other, the differences are motivated by phonotactic constraints of the recipient languages. However, more significant differences may be an indication that the way of transmitting the Arabic phrases to African languages went through other African languages, for example, the expression *Allāh akbar* ‘God is the Greatest!’ is in Hausa *Allāhū àkbar*, which is a close counterpart of the original form, whereas in Kanuri its equivalent is *álakuwar*, in Maninka *álakìbaro*. In Wolof, the expression *astafurlaa*

⁶ Following Bichler (2000: 231).

⁷ The restriction is mentioned in a course-book for spoken Hausa (Cowan & Schuh 1976: 27).

⁸ Saying *Alhamdù lillāhī* is termed as *hamdalā*.

‘God forgive me!’ has reference to an original Arabic phrase *’istağ far ’allāh* ‘I ask God’s forgiveness!’ (Baldi 2003: 7). Some other Arabic words have similar distribution in the whole area. The terms *ʔal-ājj* ‘one who has done the pilgrimage’, *ʔal-āfiya* ‘(good) health’ function in a number of phonological variants across West African countries (Yoruba *àlájì* ‘Muslim pilgrim to Mecca’, *àláfìà/àláfìyà* ‘good health’ respectively). Moreover, some religious phrases have extended the basic meaning in particular languages, e.g. Ar. *ʔistafar ʔallāh* ‘I ask God’s forgiveness!’ is in Yoruba *sáàfùlà* ‘Respected Sir!’ (Baldi 2001: 45ff).

3. The equivalents of the Arabic phrases in African languages

Religious phrases of Arabic origin have their equivalents in African languages. They are either direct translations or extension of the translated phrase into the whole set of similar phrases. The expression *Alhamdu lillahi* functions in Hausa in the original form (‘thanks be to Allah’) but also as its translated version with a personal pronoun *Mun gode wa Allah* ‘we thank Allah’. In Wolof, *alhamdulillah* and *ñungi sant yalla* are direct equivalents (Bichler 2000: 231) that are used alternatively or even in a direct sequence of two language variants.

There is a number of prepositional phrases containing the word Allah which may be referred to Arabic expression *wa-llāh* ‘by Allah’ which functions as an interjection to increase the credibility of the statement. In Hausa, the form *don Allah* ‘for God’s sake’ (lit. for Allah) expresses the request to do something or to give a swift response to the request. In Bambara, there are postpositional phrases *Ala kama*, *Ala koson*, *Ala ye* (cf. Konaré 1998: 120, 76) that have similar functions.

Languages have developed phrases that are stylistic means to manifest cultural values. The word Allah is an obligatory constituent of these expressions. The Kanuri phrase *Alada ajabba* is an equivalent of Hausa *Allah mai iko* (lit. Allah being an authority), both are translated versions (Umarbe 2014 and Umar 2014 respectively) of the relevant phrase from *Amina* (Umar 2005: 5) ‘How wonderful Allah can be’ which is used as a concluding statement of positive evaluation.

The examples show that Allah expressions in African languages are basic items of language lexicon and phraseology that secure proper communication according to the norms of the language. For that reasons they are usually introduced into the teaching materials for the West African languages as their basic elements. For example, the course-book *Je parle bien bamanan* (Konaré 1998) includes the phrases *A ka dɔgɔn Ala ye!* ‘We have to rely on Allah’, *Ala ka kenya ke i ye* ‘May Allah give him health’ that are essential for the speakers of the Bambara language.

4. “Allah expressions” in Hausa

In the following sections, the questions related to the status of religious phrases in West African languages will be illustrated by the examples extracted from one language, namely Hausa. The sociolinguistic status of this language allows tracing the use of particular

expressions in different types of texts which represent non-religious discourse, such as everyday communication, literature or specialized language communication.

4.1. Religious phrases as communication rituals

Religious phrases are linguistic means of expressing the cultural code in which the language functions. Particular phrases (either in their original Arabic forms or in translated versions) require the appropriate occasion for their use. Here are the examples from Hausa that show the variety of conventionalized phrases of religious connotation that have various communicative functions, the symbolic function in particular. Very often the Arabic expressions used as greeting formulas are combined with the Hausa equivalents. They are not interpreted as religious statements, but rather as petrified phrases that in particular context have a concrete meaning. The following example presents the situation when Nuhu comes to visit Hasan at his compound and the greetings of Arabic origin (*Salamu alaikum* and *amin*, *alaika salamu* in response) are used with the alternative Hausa greeting (*Barka da zuwa* ‘greetings on your coming’, ‘you are welcome’), i.e. (Cowan & Schuh 1976: 35):

(9) Nūhù: *Sàlāmù àlaikùm*. ‘May I come in?’

Hasàn: *Āmin, àlaikà sàlāmù. Barkà dà zuwà*. ‘Come on in. Greetings on your coming.’

The expression *Alhamdulillah* functions in Hausa in the original form (‘thanks be to Allah’) but also as its translated version with a personal pronoun *Mun gode wa Allah* ‘we thank Allah’. As a response to a concrete question, the two versions may be used, as for example while asking the farmer about his crops (Yahaya & Yusufu 1989: 36):

(10) *Daminar bana kam alhamdu lillahi. Mun gode wa Allah.*

‘As for rainy season of this year, *alhamdu lillahi*. Thanks be to Allah’

The tautologic sequence of the two phrases in two different languages is a strategy to strengthen the message (i.e. positive response) in terms of the cultural values.

Most of the religious phrases invoking the name of Allah function as communication rituals which should be used in a proper context. For example, when we hear the news on a newly born baby, the proper reaction is *Allah ya raya* ‘may Allah give him/her (child) long life! (lit. ‘may Allah cause him to grow up’). When it is about someone’s death, it is to say *Allah ya ji kansa, ya yi masa rahama* ‘may Allah have mercy on his soul!’

There are a number of phrases used to say goodbye that involve the lexeme Allah. The following variants of farewell statements are in fact good wishes that include the request to Allah for their fulfilment:⁹

(11) a. *Allah ya kai mu* ‘may we live to do it as you wish!’

b. *Allah ya bar mu tare* ‘may Allah leave us together’

⁹ The examples of statements suitable for various occasions are presented in pedagogical books, e.g. (Bature et al. 1998; Cowan & Schuh 1976).

- c. *Allah ya ba mu lafiya* ‘May Allah give us health’¹⁰
- d. *Allah ya ja zamaninka* ‘long life to you!’
- e. *Allah ya ba ka nasara* (said to senior) ‘if I might venture to make a suggestion...’

Therefore, ‘thank you’, out of *na gode* has the extension of *Allah ya saka da alheri* ‘may Allah give a good reward!’, *Alla ya ba da lada* ‘may Allah reward you!’

The phrases invoking Allah in Hausa are not so specific for their construction or stylistic variations, but rather for their rooting in the language, which is manifested in communication practice of the spoken language. Their significance is confirmed in written texts of various discursive properties.

A special type of expressions is characteristic mainly of the written texts. Whenever the name Allah appears, the abbreviation S.W.T is used. This is an acronym for the Arabic words *subhanahu wa ta’ala* which means ‘The most glorified, the most high’. A Hausa equivalent may be added, i.e. *Mai Girma da Daukaka*, but the abbreviation (S.W.T.) is always left in original form. Another abbreviation is used after the mentioning of the Prophet Muhammad, namely (S.A.W. or saws) which corresponds to the Arabic phrase *sallahu alayhi wa sallam* ‘May Allah bless him and grant him peace’, *tsira da amincin Allah su tabbata gare shi* in Hausa version. The phrases and the justification for their use are taken directly from the religious treatises, but the pattern is spreading and functions also in other texts, including the academic publications dealing with language or literature.

The religious phrases of Arabic origin and their Hausa counterparts that are inserted in the main text have predominantly symbolic functions. They are used to manifest cultural identity which has Islamic foundations.

4.2. The name of Allah in phraseology and discourse

Hausa is seen as a figurative language in which metaphoric expressions are highly appreciated with regard to stylistic convention. Phraseology involving Allah includes expressions that are used as metaphors in which the notion (action, event, state) is conceptualized as a result of what was ordered by Allah. Therefore, the meaning ‘they were married’ is stylistically more preferably expressed as *Allah ya nufe su da yin aure*, i.e. ‘Allah destined them to be married’. Similarly *Allah ya kaddara shi da haihuwa* ‘he has got a baby’ in literal translation is ‘Allah ordained that the child should be delivered to him’. This conceptual pattern is often used to express states (such as tiredness), e.g. *Allah ya gajishe shi* ‘he became tired’ (lit. Allah made him tired). The above examples were extracted from the novel *Magana Jari Ce* to confirm that the beauty of language is also determined by the use of the phraseology of religious connotation.¹¹ Here are also some other examples taken from the novel *Idan So Cuta Ne...* [When love is a disease] by Y. M. Adamu. The novel printed on 114 pages includes 91 phrases distinguished by the use of the word Allah. It should be added that this is a novel about love, the main characters are young people

¹⁰ The dictionary (Abraham 1962) gives the extension of this meaning into: ‘well, we’ll wait and see what happens!’

¹¹ The metaphoric expressions involving the name of Allah are occasionally used in scientific texts, for example the statement ‘there are three sounds representing the phoneme /f/’ has been rendered as “Allah endowed the language with three sounds as a representation of the phoneme /f/” (own notes).

and the dialogs reflects every-day conversation. Out of ritual communication: greetings, condolences, wishes, the phrases with the name Allah have many discursive functions, e.g.:

- expressing hope for a desired outcome
- (12) *Yaya jikin Park din? – Da sauki, da yardar Allah zai sami lafiya ba da dadewa ba.* (p. 77)
‘How is Park? – He is recovering, **we hope that** (i.e. with the permission of Allah) he will be well soon’
- justifying the current situation with the will of Allah
- (13) *Allah ya sa Hajiya na nan kusa* (p. 96)
‘By chance (i.e. **Allah caused**) Hajiya was around’
- expressing emotions
- (14) *Wallahi tallahi Farida, ina kaunarki fiye da kowa* (p. 60)
‘**(Great Allah!)** Farida, I love you so much!’
- confirmation, agreement
- (15) *Haba? Farida ta tambaya* (p. 16) Really? Farida asked.
Allah kuwa. Hadiza ta ba da amsa ‘**Of course** (i.e. Allah indeed). Hadiza answered.’

The text of the novel confirms the ubiquitous use of the religious phrases in non-religious contexts. It makes them not recognizable as religious phrases, but the most appropriate systemic means to render the meaning.

5. Desacralization of religious phrases

The increasing usage of religious phraseology in various texts is linked to a reverse process which is connected with the loss of their religious connotation. Many religious phrases, the so-called “Allah expressions” in particular, are lexical units or systemic devices to code grammatical relations.

In Hausa, the word derived from *Alla* (‘Allah’) is an instance of conversational routine that has changed the status of the noun. When pronounced with interrogative intonation (*alla?*) it carries the meaning ‘really?’ The reply *alla* means ‘yes, truly’.

The contexts in which the Arabic phrases are used, give them new syntactic attributes. As already mentioned, the phrase *in sha Allahu* (lit. ‘if Allah wills’) may be a lexical equivalent of ‘future’, ‘(to) hope’ or ‘fine’, depending on the context. Similarly *Madalla*, which is etymologically related to the Arabic phrase meaning ‘thanks be to Allah’¹², has

¹² Following Abraham (1962: 629).

gained new pragmatic functions. In conversation it is used to acknowledge the acceptance of what was said (‘Fine!’ Good idea!) and to close the discussion or change the topic to another one, e.g. *To, madalla. Sai mu tafi.* ‘Fine, Let’s go’.

Religious phrases of Arabic origin are commonly used in Hausa as interjections which show emotions (sadness, excitement, among others). For example *Allahu akbar!* ‘God is the Greatest! is equal to ‘how sad!, whereas *Allah sarki!* (lit. God is the Lord) to ‘how wonderful!’

Some phrases have undergone the process of semantic and functional modifications in relations to the original Arabic source. The Hausa word *bismilla* is referred to the Arabic phrase *Bismillahi-rrahmani-rrahim*. بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ ‘In the name of Allah, the most gracious the most merciful’. Many Muslims use a full phrase before reciting the Qur’an. In colloquial Hausa, *bismilla* is a formulaic invitation to begin a meal, come into a room, sit down, and so forth. It is the word that maintains communication rituals and belongs to the sphere of politeness. In modern communication, it is a secular rather than religious phrase.

Similarly, some other phrases have the status of lexical units rather than interjections or any other metatextual operators. The phrase *don Allah* is basically an interjection (‘for God’s sake’) which expresses the request, pressure or insisting on doing something. It may be used in dialogs with the meaning ‘please (do it)’, but also in indirect (reported) speech when it is a phrase modifier, e.g. (examples from *Idan So Cuta Ne...*):

(16) *Ta nuna masa gidansu ta ce don Allah ya zo*
‘she has shown him her house and **insisted** that he should come’

(17) *Don Allah ina son magana da kai* (p. 77)
‘I **really** want to speak to you’

The independent status of the unit *allā-allā* ‘eagerness’ and the phrase *inā yā Allā, bābū yā Allā* (lit. where is Allah, there is no Allah) is noted in the dictionary (Abraham 1962: 25f). The lexical meaning of these petrified forms is confirmed in a novel by Abubakar Imam (*Magana Jari Ce*, p. 6), i.e.:

(18) *Yana alla-alla Sarki ya mutu bai haihu ba*
‘he was **overwhelmed** (with thought of what may happen) when the Emir dies childless’

(19) *Ana nan ran nan, ina ya Allah babu ya Allah sai Waziri ya sami wata dabara*
‘Wizier **suddenly** bethought himself of a plan’

Desacralization of religious phrases is an advanced process in Hausa, but it is also the ongoing process in many other African languages, the detailed exemplification, however, needs further research.

6. Summary

The religious phrases, along with a number of lexical borrowings introduced with Islam into West African languages have been incorporated into the structures of particular languages and developed as their independent items. They are commonly used in ritual communication (greetings, exclamations, etc.), but they also enrich systemic devices of expressing hope, desire, wishes and beliefs. As they are used not only in religious context but also in everyday communication, their status is changing. Arabic expressions in West African languages, similarly to the Bible phraseology in European languages in the seventeenth century (de Grazia 1980: 319), are involved in the processes of secularization.

From the linguistic perspective, West Africa is a big area of convergence in which religious phrases of Arabic origin function. They are recognizable throughout the area for their phonological similarity, but mostly for the same cultural code which they represent. Arabic religious phrases, along with other Arabic borrowings, have penetrated regions in which Islamic religion is not dominant as far as the Yoruba society in the south.

The religious vocabulary and religious phrases of Arabic origin are part of the inventory that provides convincing arguments of identifying common cultural area in West Africa, similarly to European cultural area which is distinguished on the basis of Latin vocabulary and the Bible phraseology common to these languages (Nagórko 2011; Burkhardt et al. 2012).

Bibliography

- Abraham, Roy Clive. 1962. *Dictionary of the Hausa language*, London: University of London Press Ltd.
- Adamu, Yusuf Muhammad. 1989. *Idan So Cuta Ne...* Kano: A.J. Publishers.
- Baldi, Sergio. 2008. *Dictionnaire des emprunts arabes dans les langues de l'Afrique de l'Ouest et en swahili*. Paris: KARTHALA.
- Baldi, Sergio. 2001. "On Arabic Loans in Yoruba". In: Dymitr Ibrizimow, Rudolf Leger & Uwe Seibert (eds.). *Von Ägypten zum Tschadsee. Eine linguistische Reise durch Afrika*. Festschriften für Herrmann Jungraithmayr zum 65. Geburtstag. (Abhandlungen für die Kunde des Morgenlandes, Band LIII, 3). Würzburg: Ergon-Verlag, 45-65.
- Baldi, Sergio. 2003. "On Arabic Loans in Wolof: Addenda". *Studi Magrebini* 1 (Nuova Serie), 1-22.
- Baldi, Sergio. 2004. "Arabic Loans in Bidiya". *Studi Magrebini* 2 (Nuova Serie), 71-87.
- Bature, Abdullahi, & Russell G. Schuh. 1998. *Hausar Baka – "Gani Ya Kori Ji"*. Windsor, California: World of Languages.
- Bichler, Gabriele Aïsha. 2000. *Bejo, Curay Und Bin-Bim? Eine soziolinguistische Studie über den Zusammenhang zwischen Spracherwerb und Kulturkenntnis am Beispiel der Wolof in Dakar, Senegal*. Ph.D. dissertation, manuscript, Universität zu Wien.
- Burkhardt, Hanna & Robert Hammel, Marek Łaziński (eds.). 2012. *Sprache im Kulturkontext: Festschrift für Alicja Nagórko*. Frankfurt am Main: Peter Lang.
- Clift, Rebecca. & Fadi Helani. 2010. "Inshallah: Religious invocations in Arabic topic transition". *Language in Society* 39, 357–382.
- Cowan, J. Ronayne & Russell G. Schuh. 1976. *Spoken Hausa*. Ithaca, New York: Spoken Language Service, Inc.
- Greenberg, Joseph. H. 1947. "Arabic Loan-Words in Hausa". *Word* 3, 85-97.
- De Grazia, Margreta. 1980. "The Secularization of Language in the Seventeenth Century". *Journal of the History of Ideas* 41(2), 319-329.

- Imam, Abubakar. [1937] 1980. *Magana Jari Ce*. Zaria: NNPC.
- Konaré, Demba. 1998. *Je parle bien bamanan*. Bamako.
- Löhr, Doris & Ari Awagana. 2009. “Loanwords in Hausa. Results from the Loanword typology project”. In: Eva Rothmaler (ed.). *Topics in Chadic Linguistics. Band V. Papers from the 4th Biennial International Colloquium on the Chadic Languages, Bayreuth, October 30-31 2007*. Köln : Rüdiger Köppe Verlag, 37-49.
- Nagórko, Alicja (ed.). 2011. *Sprachliche Säkularisierung (Westslawisch – Deutsch)*. Hildesheim: Olms Verlag.
- Umar, Mohammed Kabir. 2005. *Amina* (English version). Africa World Press.
- Umar, Mohammed Kabir. 2014. *Amina* (Hausa version). Africa World Press.
- Umarbe, Mohammed. 2014. *Amina* (Kanuri version). Africa World Press.
- Yahaya, Ibrahim Yaro & Yunusa Yusufu. 1989. *Da koyo akan iya, Littafi na uku*. Ibadan University Press PLC.
- Yalwa, L. Danladi. 1992. “Arabic Loan Words in Hausa”. *Ufahamu: A Journal of African Studies* 20(3) 100-131.

The monograph covers the main aspects of studies on West African languages related to the diversity of structural patterns and complexity of their linguistic assignment. It includes various topics ranging from linguistic description and conceptualization patterns to the sociolinguistics of contemporary refugee camps. Typological diversity is enriched with the presentation of pidgin structures and sign languages. Structural differences between languages are seen from a comparative perspective, which also indicates the areal dimension of linguistic processes. The presentations of linguists from both Europe and Africa develop the idea of convergence area in West Africa, which is motivated by the contact between languages of different affiliations to language families and common cultural basis of language development.

Nina Pawlak is a professor of African linguistics, member of the Department of African Languages and Cultures, University of Warsaw. Her academic interests focus on Hausa, in comparative perspective they also cover other Chadic languages. The main fields of linguistic investigation are typology, language development, cognitive linguistics, culture and communication.

Izabela Will is an assistant professor at the Department of African Languages and Cultures at the University of Warsaw where she teaches Hausa, general linguistics, culture of Nigeria and African art. Her academic work focuses on Hausa language, culture of Northern Nigeria, and gestures produced by speakers of Hausa.