

# Clause structure, pro-drop and control in Wolof: an LFG/XLE perspective\*

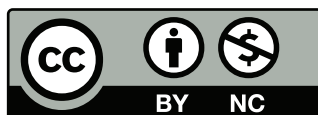
*Cheikh Bamba Dione*

*Postdoctoral Researcher  
Department of Linguistics  
University of Bergen  
dione.bamba@uib.no*

## **Abstract**

This paper provides a formal description of the syntactic analysis of core constructions of Wolof clausal/verbal morphosyntax within the Lexical-Functional Grammar formalism. This includes the basic phrase structure, pro-drop, and control relations. The Wolof grammar is implemented in XLE and uses a cascade of finite-state transducers for morphological analysis and tokenization. This work is part of the ongoing process on building language resources and tools for Wolof, in particular a computational grammar.

**Keywords:** computational grammar, control, LFG, morphosyntax, pro-drop, Wolof



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

---

\*I would like to thank the editor and the anonymous reviewers of of NJAS for helpful comments. I also want to thank collaborators from the Infrastructure for the Exploration of Syntax and Semantics (INESS) infrastructure and from the ParGram project.

## **About the author**

Cheikh M. Bamba Dione is a postdoc researcher in linguistics at the Department of Linguistic, Literary and Aesthetic Studies of the University of Bergen. His interests include theoretical and computational linguistics as well as psycholinguistics. He has done extensive work on Wolof morphology and grammar. His most recent research focused on the implementation of computational models to develop digital language resources. He is also conducting studies on the use of statistical and machine learning approaches to develop language technology applications.

## 1 Introduction

This paper presents a Lexical-Functional Grammar (LFG) (Kaplan and Bresnan, 1982) description of Wolof, a Niger-Congo language mainly spoken in Senegal, Gambia and Mauritania. In particular, it discusses LFG-based analyses proposed for Wolof to model the basic phrase structure, pro-drop, and control relations in that language. This research work has taken place within the Parallel Grammar (ParGram) project (Butt et al., 1999b; Butt et al., 2002) which is itself embedded within LFG. The aim of ParGram is to produce deep, linguistically well-motivated, and maximally parallelized grammars for a variety of languages (Sulger et al., 2013). The ParGram grammars are developed using the Xerox Linguistic Environment (XLE) (Crouch et al., 2019), an efficient grammar development platform.

The paper is organized as follows. Section 2 gives a short overview of LFG and briefly discusses LFG parsing with the XLE platform. Section 3 provides background information on Wolof relevant for the present discussion. Section 4 examines various Wolof constructions in an LFG setting. Finally, section 5 concludes the discussion.

## 2 Overview of LFG / XLE

Traditional LFG analyses focus on two levels of syntactic representation: constituent structure (c-structure) and functional structure (f-structure) (Kaplan and Bresnan, 1982). C-structure is represented as a phrase structure tree and models the surface exponence of syntactic information, e.g. precedence, dominance, constituency and syntactic categories. F-structure includes functional syntactic concepts, i.e. grammatical functions (GF) such as subject, object, predicate and grammatical features, e.g. person, number. F-structures are represented as feature structures or attribute-value matrices (AVM) (Butt et al., 1999a). While c-structure may vary widely between languages, f-structure information remains relatively constant across languages. This follows LFG's universality principle which assumes that "internal structures are largely invariant across languages" (Bresnan, 2001, page 45).

By way of illustration, let us consider the Wolof sentence in (1).<sup>1</sup>

- (1) *Móodu lekk na jën*  
 Móodu eat +F.3sg fish  
 "Móodu has eaten fish"

This sentence can be analyzed in LFG using the phrase structure rules in (2).

- (2) a.  $IP \rightarrow \begin{matrix} NAMEP & S \\ (\uparrow SUBJ)=\downarrow & \uparrow=\downarrow \end{matrix}$       b.  $S \rightarrow \begin{matrix} VP \\ \uparrow=\downarrow \end{matrix}$   
 c.  $VP \rightarrow \begin{matrix} V' & NP \\ \uparrow=\downarrow & (\uparrow OBJ)=\downarrow \end{matrix}$       d.  $NAMEP \rightarrow \begin{matrix} NAME \\ \uparrow=\downarrow \end{matrix}$   
 e.  $V' \rightarrow \begin{matrix} V & Cl \\ \uparrow=\downarrow & \uparrow=\downarrow \end{matrix}$       f.  $NP \rightarrow \begin{matrix} N \\ \uparrow=\downarrow \end{matrix}$

<sup>1</sup>Abbreviations in the glosses: Cinf: infinitival complementizer; cl: noun class; CLF: classifier; COMP: complementizer; DET: determiner; DFP: definite proximal; DFD: definite distal; +F: finite; FUT: future; INF: infinitive; IMP: imperative; IPF: imperfective; NDF: indefinite article; NEG: negation; NSFOC: non-subject focus; OPT: optative; PRONexpl: expletive pronoun; PRES: present; PST: past; pl: plural; S: subj; sg: singular; SFOC: subject focus; VFOC: verb focus; 1, 2, 3: first, second, third person.

The XLE parser coupled with these grammar rules assigns to the sentence in (1) the c- and f-structure given in Figure 1

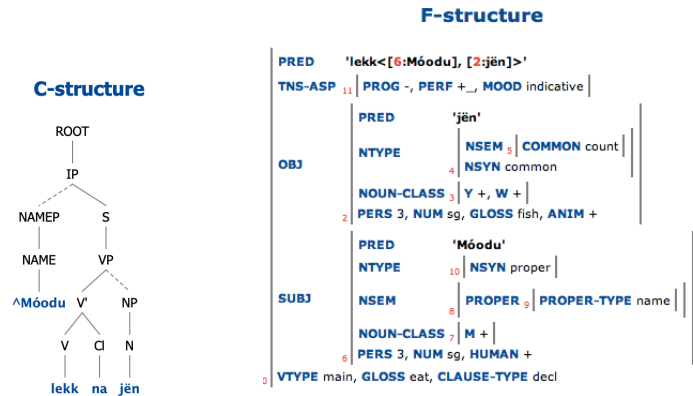


Figure 1: C- and f-structure of sentence (1)

The two structures in Figure 1<sup>2</sup> are related to each other by a functional projection function  $\phi$  (Kaplan and Bresnan, 1982) from c-structure nodes to f-structure AVMs. This relationship is explicitly stated on the phrase structure rules of the LFG grammar (2). Figure 2 shows the relationship between the c-structure and the f-structure of (1), with the  $\phi$  projection indicated by arrows leading from phrase-structure nodes to AVMs.<sup>3</sup>

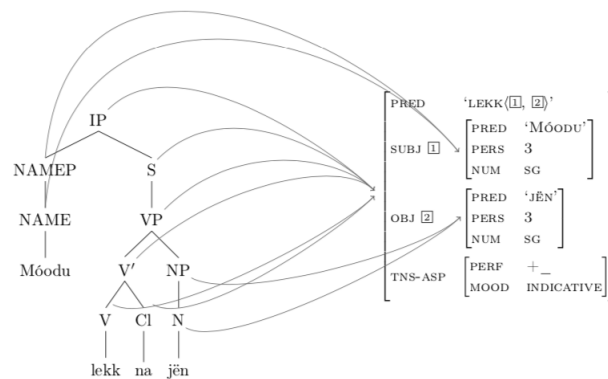


Figure 2: The correspondence between c-structure and f-structures

The relationship between c-structures and f-structures are encoded through so-called functional equations (Kaplan and Bresnan, 1982). The metavariables  $\uparrow$  and  $\downarrow$  respectively refer to the mother c-structure node and the non-terminal node itself. Grammatical information expressed in these equations may also come from the lexicon. For instance, the grammatical representation for (1) includes information about grammatical features (e.g. number, person, tense, aspect) that is encoded in the lexicon, as shown in (3).

- (3) *Móodu* NAME ( $\uparrow$  PRED)='Móodu'  
 ( $\uparrow$  PERS)=3.  
*lekk* V ( $\uparrow$  PRED)='lekk<( $\uparrow$  SUBJ)( $\uparrow$  OBJ)>'.  
*na* Cl ( $\uparrow$  PRED)='na'  
*jën* N ( $\uparrow$  PRED)='jën'

<sup>2</sup>This output was done via XLE-Web (part of INESS, <http://clarino.uib.no>). See Dione (2013) for the c-structure analysis proposed for Wolof (e.g. the analysis of *na* as a clitic (Cl)).

<sup>3</sup>For ease of exposition, some features and values in this f-structure have been omitted.

<i>na</i>	CI	(↑ TNS-ASP PERF)=+_ (↑ SUBJ NUM)=sg (↑ SUBJ PERS)=3.
<i>jën</i>	N	(↑ PRED)='jën' (↑ PERS)=3.

In (3), the constraint (↑ PRED)='Móodu' states that the preterminal node immediately dominating the terminal symbol *Móodu* has an f-structure whose value for the attribute PRED is 'Móodu'. The entry also contains information as to person, which is relevant for subject-verb agreement. The entry *lekk* is analyzed as a verb (V) that subcategorizes for a SUBJ and an OBJ. The inflectional element *na* is treated here as a clitic (CI) (see Dione, 2013) that expresses a combination of features, some subject-related (i.e. subject agreement) and some clausal (e.g. the perfective aspect of the clause).

The examples given above show simplified full-form lexicon entries. But in fact, as most of the ParGram grammars, the Wolof grammar uses a cascade of finite-state transducers (FST) (Kaplan et al., 2004) to preprocess the input. The first FST acts as a tokenizer and a normalizer (Dione, 2017) which splits the input stream into a unique sequence of tokens separated by whitespaces (e.g. space, line break) or by punctuation characters. When surface strings are looked up, the output string is the input string plus a special symbol, i.e. TB (for "Token Boundary") inserted between the tokens. For instance, the sentence in (1) is tokenized as shown in (4). Because this sentence only contains words that are clearly separated by whitespaces, tokenization is quite straightforward. However, in many other cases, tokenization faces non-trivial issues related to multi-word expressions (MWEs), clitics and text normalization, which require language-specific information.<sup>4</sup>

(4) Móodu lekk na jën. ⇒ Móodu TB lekk TB na TB jën TB

Next, the input is preprocessed by the Wolof Morphological Analyzer (WoMA) (Dione, 2012). With this tool, morphophonological properties of words are modeled within a finite-state transducer and interfaced with the syntax. Based on the Xerox finite-state tool, *fst* (Beesley and Karttunen, 2003), WoMA handles the input in both directions: analysis and generation. A surface form is associated with a canonical form (stem or lemma) and a set of tags encoding morphological features.

- (5) a. Móodu ⇔ Móodu+Name+PropNoun+PropTypeName+Human  
 b. lekk ⇔ lekk+Noun+Common+b+y  
 ⇔ lekk+Verb+Base+Main+Active  
 c. jën ⇔ jën+Noun+Common+b+y

Example (5a) states that the surface form *Móodu* can be analyzed as a stem *Móodu* and a proper noun (+Name). The feature +PropNoun and +PropTypeName respectively specify the syntactic and semantic type of the lexical entry. The former distinguishes between common nouns, proper nouns and pronouns. The latter differentiates common nouns from proper nouns. Semantic features of the nouns are usually useful in constraining syntactic constructions. +Human indicates that the surface form is associated with humanness features.

<sup>4</sup>See Dione (2017) for more details.

Example (5b) illustrates cases where a single form corresponds to more than one analysis. The surface form *lekk* can be analysed as a common noun (i.e. ‘food’) that agrees with its modifier in the *b* and *y* noun classes (see section 2.1) or as the base form of an active verb (i.e. ‘eat’). Example (5c) is analyzed in a way that is similar to the first reading of the entry *lekk*.

Having briefly presented an overview of LFG/XLE and some relevant components of the Wolof grammar, let us now consider the morphosyntax of this language.

## 2.1 Wolof nominal system

### 2.1.1 Nouns and noun classes

Wolof is an SVO language with a complex noun class agreement system (McLaughlin, 1997; Tamba et al., 2012). The language has 8 singular and 2 plural noun classes. These are identified by their index, which functions as a stem to which a determiner/pronoun affix is added. Singular classes are: *b*, *g*, *j*, *k*, *l*, *m*, *s*, *w*. Plural classes are *y* and *ñ*. As for plural noun classes, *y* is the class of most nouns, while *ñ* is the class of a restricted small set of human nouns. Unlike the noun class system found in Bantu languages, nouns in Wolof lack a class marker on the noun itself. Instead, class membership is marked on the determiners, e.g. definite and indefinite articles, demonstratives, relative pronouns.

### 2.1.2 DP structure

The determiner system of Wolof is built around three determiner vowels *u/i/a* and a numeral-like expression (Tamba et al., 2012). The language possesses two definite and two indefinite articles, all agreeing in class with the noun phrase (NP). However, indefinite and definite determiner phrases (DPs) have a different word order. While the definite article obligatorily follows the NP (6-7), the indefinite article obligatorily precedes the NP (8). Thus, the language displays typologically mixed head-initial/head-final characteristics, e.g. post-nominal definite determiners, but pre-nominal indefinite determiners.

- |   |  |  |
|---|--|--|
| (6) <i>xale b-i</i><br>child cl-DFP<br>“The child (here)” | (7) <i>xale b-a</i><br>child cl-DFD<br>“The child (there)” | (8) <i>a-b xale</i><br>NDF-cl child<br>“A child” |
|---|--|--|

The vowel suffixes *i* and *a* in the definite articles in (6) and (7) respectively encode proximity and distance in space, time, or conversation (Torrence, 2005). In contrast, the vowel prefix *a* in (8) marks indefiniteness. The Wolof articles typically consist of a class index glossed as *cl* and one of the determiner vowels *u/i/a*. Accordingly, proximal, distal definite articles and the simple indefinite article have the patterns *cl-i*, *cl-a* and *a-cl*.

## 3 Wolof clausal/verbal morphosyntax

Two main aspects of the Wolof verbal system have been mostly studied in the literature (Robert, 1991; Ka, 1994). First, the main verb usually does not itself carry inflectional markers. Second, Wolof exhibits a complex system of inflectional elements, pronouns or clitics that appear as separate words or as verbal suffixes. The inflectional markers can express subject agreement, but also aspect, polarity, and the focus in the sentence, leading to interesting interactions between these elements. Accordingly, several different paradigms of the inflectional markers are available. These vary in both form and linear position depending on the kind of information

they express in the sentence (Torrence, 2003). As a result, Wolof exhibits a number of different clause types based on these paradigms.

To provide the reader with a general understanding of these inflectional markers, section 3.1 will first discuss the typology of Wolof clauses. Then, section 3.2 will examine the finiteness status of Wolof verbs.

### 3.1 Wolof clauses

The clausal organization of Wolof is built around various clause types (see Zribi-Hertz and Diagne, 2002; Torrence, 2005). This includes affirmative clauses, negative clauses, subject focus, non-subject focus, verb focus clauses, optative clauses, progressive clauses, narrative clauses and non-finite verbal complement clauses.

Let us begin with independent affirmative clauses, as illustrated in (9-10). In these examples, the entire clause is new information, and no constituent is in focus. The perfective aspect is expressed via a combination of the verb (*lekk* ‘eat’) with the subject marker *na-ñu* which also carries person and number information.<sup>5</sup>

(9) *Xale yi lekk na-ñu ceeb*  
 child the.pl eat +F-3pl rice  
 “The children have eaten rice.”

(10) *Lekk na-ñu ceeb*  
 eat +F-3pl rice  
 “They have eaten rice.”

The contrast between (9) and (10) exemplifies the pro-drop nature of Wolof. The constituent *xale yi* in (9) is optional, meaning that the sentence can lack an overt subject, and remains nevertheless grammatical. This phenomenon is known to occur across a wide range of languages, including Chicheŵa (Bresnan and Mchombo, 1987; Butt, 2007). For instance, the sentence in (11a) has both an NP subject (i.e. *njûchi*) and an agreement marker (10.SM, where SM stands for subject marker). In contrast, in (11b), there is no subject NP. In this case, the subject marker supplies the referential argument that satisfies the argument slot of the verb.

(11) a. *Njûchi zi-ná-lúm-a a-lenje* (*Chicheŵa*)  
 10.bee 10.SM-Past-bite-FV 2-hunter  
 “The bees bite the hunter.”

b. *Zi-ná-lúm-a a-lenje* (*Chicheŵa*)  
 10.SM-Past-bite-FV 2-hunter  
 “They bit the hunters.”

In many languages (including Wolof and Chicheŵa), pro-drop correlates with pronoun incorporation or pronominal inflection (Bresnan and Mchombo, 1987; Bresnan, 2001). This means that the core arguments of the verb are not obligatorily expressed in terms of syntactic constituents projected from independent morphological words such as personal pronouns. If an argument is

<sup>5</sup>In the literature, the clauses in (9-10) are referred to as *na* clauses (Tamba et al., 2012), perfective clauses (Robert, 1991) and ‘no-focus’ clauses (Ka, 1994). I follow the analysis of *na* proposed by Zribi-Hertz and Diagne (2002) as finiteness markers, accordingly glossed as +F. However, I generally assume finiteness to be indicated by a combination of the lexical verb with the *na* marker, which serves as a host for the clitic (e.g. *-ñu*) marking person and number information in the clause.

realized, the person/number markers in the verb function as agreement features. Otherwise, if the argument is missing, then the markers provide a pronominal interpretation. Thus, in (10) and in (11b), the overt subject can be missing, because the language (e.g. Wolof and Chicheŵa, respectively) freely allows the omission of such an argument. However, there are many other pro-drop languages such as Urdu (Butt, 2007) and Japanese (Masuichi et al., 2003) in which the correlation between pro-drop and agreement does not seem to hold. In other words, such languages allow arguments to be dropped without any corresponding agreement morphology (see section 4.2).

To come back to the description of the Wolof clause types, the clause in (12) represents the negation of (10). Clausal negation is expressed by the inflectional morpheme *-u*. As with examples (9-10), in (12), the entire clause expresses new information, meaning that no constituent is in focus. Again, (12) illustrates a sentence with a dropped NP subject.

- (12) *Lekk-u-ñu ceeb*  
 eat-NEG-3pl rice  
 “They did not eat rice.”

In contrast to (9-10) and (12), examples (13-15) illustrate sentences which have their information structure explicitly marked in their morphosyntax. Wolof has grammaticalized, morphosyntactic means of expressing focus. The language distinguishes three kinds of focus (Church, 1981; Robert, 1991; Torrence, 2005): subject focus, non-subject focus and verb/predicate focus. Clefting can be used to put the subject (13), the predicate (14), and any constituent which is neither subject nor main verb (15) into focus.

- (13) *Xale yi ñoo lekk ceeb* (Subject focus)  
 child the.pl SFOC.3pl eat rice  
 “It the children who have eaten rice.”
- (14) *Xale yi da-ñu lekk ceeb* (Verb focus)  
 child the.pl VFOC-3pl eat rice  
 “The children did eat rice.”  
 “Eat rice is what the children did.”
- (15) *Ceeb la xale yi lekk* (Non-subject focus)  
 rice NSFOC.3 child the.pl eat  
 “It’s rice that the children ate.”

Note that in non-subject focus clauses, if the lexical *DP* subject is overtly realized as in (15), person inflection (e.g. *ñu*) does not attach to the focus morpheme *la*. This yields the surface form *la* instead of *la-ñu*. The form *la-ñu* is only allowed when the sentence does not have an explicit lexical subject as in (16). Otherwise, a co-occurrence of the lexical *DP* subject and the subject agreement marker *ñu* causes ungrammaticality (17).



(16) *Ceeb la-ñu lekk*  
 rice NSFOC-3pl eat  
 “It’s rice they ate.”

(17) \**Ceeb la-ñu xale yi lekk*  
 rice NSFOC-3pl child the.pl eat

In Wolof, focus is marked morphosyntactically. The focus marker takes a different form depending on the focus type, the person and number of the subject. In addition, there is a variation in word order depending on the focus type. The focus marker precedes the focused constituent in verb focus clauses (14), but follows it in both subject (13) and non-subject (15) focus clauses.

Example (18) illustrates optative clauses which are used to express wishes or desires of a speaker.<sup>6</sup>

(18) *Na-ñu lekk ceeb*  
 Opt-3pl eat rice  
 “May they eat rice.”

(19) *Lekk na-ñu ceeb*  
 eat +F-3pl rice  
 “They have eaten rice.”

Optative mood is expressed by means of both morphology and word order. In optative clauses (18), the subject marker (e.g. *nañu*) precedes the lexical verb. In contrast, in finite perfective clauses (e.g. (10), which is repeated in (19)), the subject marker follows the verb. Likewise, the difference in terms of morphology can be seen by considering the second person singular. While in the third person plural, the finite (19) and the optative (18) forms are the same, these forms overtly differ in the second person singular: finite 2SG = *nga*; optative 2SG = *nanga*. This difference can be seen by comparing example (21) with (20). This means that the subject marker *nañu* exhibits here a case of syncretism (Baerman et al., 2005). The single form *na-ñu* in (19) and (18) corresponds to two distinct morphosyntactic descriptions (visible in the different paradigms and linear position).

(20) *Na-nga lekk ceeb*  
 Opt-2sg eat rice  
 “May you eat rice.”

(21) *Lekk nga ceeb*  
 eat +F.2sg rice  
 “You have eaten rice.”

### 3.2 Narrative, non-finite and infinitive clauses

Wolof has an intricate verb finiteness system (see Zribi-Hertz and Diagne, 2002; Zribi-Hertz and Diagne, 2003). As the Wolof clause types discussed so far illustrate, the main verb does not itself carry inflectional markers — except for a very few cases like in negative and imperative clauses. For instance, Wolof has so-called narrative clauses (22). As the name suggests, such clauses are typically found in texts with a certain type of ‘vivid’ narrative style, “formally similar to the so-called *infinitif de narration* (narrative infinitive) of French (e.g., *Et le loup de manger le chasseur* ‘So the wolf eats the hunter’” (Zribi-Hertz and Diagne, 2002, page 835). Hence, Wolof narrative clauses are typically translated by “So + PRESENT TENSE”. As for subject

<sup>6</sup>The term *optative* clauses is borrowed from Zribi-Hertz and Diagne (2002). These clauses are called *injunctive* in Ka (1994) and *obligative* in Church (1981).

agreement, unlike in the affirmative clauses discussed above, the subject marker *ñu* in (22) is a personal pronoun that precedes the verb.

- (22) *Ñu lekk ceeb*  
 3pl eat rice  
 “So, they eat rice.”

Narrative clauses express the aorist aspect (Robert, 1996). In Wolof, the aorist is unmarked and denotes a verbal notion beyond a situational anchor. It may be found in proverbs, consecutive subordinate clauses, etc., taking various temporal values, including past, present or future, according to the situational background.

Wolof narrative clauses crucially contrast with non-finite verbal complement clauses (CPs), as expressed in the bracketed constituent in (23a). Such embedded clauses may translate as English infinitivals in some contexts and typically function as complements of predicates of desire, command, wish, etc. They are similar to narrative clauses in that they are unspecified for finiteness (23b), person inflection on the verb and polarity (23c).

- (23) a. *Awa bëgg na [ñu lekk ceeb]*  
 Awa want +F.3sg 3pl eat rice  
 “Awa wants them to eat rice.”  
 b. *\*Awa bëgg na [lekk na-ñu ceeb]*  
 Awa want +F.3sg eat +F-3pl rice  
 c. *\*Awa bëgg na [lekk-u-ñu ceeb]*  
 Awa want +F.3sg eat-NEG-3pl rice

However, non-finite verbal complement clauses differ with narrative clauses in that they can be overtly marked for tense and aspect, as witnessed by (24).

- (24) *Awa bëgg-oon na [ñu d-oon lekk ceeb].*  
 Awa want-PST +F.3sg 3pl IPF-PST eat rice  
 “Awa wanted them to eat rice.”

Now, both narrative and non-finite CPs clauses should be distinguished from infinitive clauses (i.e. the bracketed constituents in (25-26)). For instance, the verbs of non-finite CPs can take a personal pronoun subject (Voisin-Nouguier, 2006), e.g. *ñu* in (23a), while the verbs in infinitive clauses cannot. In case of subordination, the latter verbs typically have an argument which shows coreference with an argument of the matrix verb. Otherwise, they have a null subject (see section 4.3).

- |  |   |
|--|---|
| <p>(25) <i>Móodu bëgg na [a lekk]</i><br/>         Móodu want +F.3sg Cinf eat<br/>         “Móodu wants to eat.”</p> | <p>(26) <i>Móodu bëgg na [lekk]</i><br/>         Móodu want +F.3sg eat<br/>         “Móodu wants to eat.”</p> |
|--|---|

Like the infinitives in English, referred to as *bare infinitive* vs. *to infinitive* (Van Valin and La Polla, 1997), Wolof infinitival clauses can surface in two different forms. The verb of the

embedded clause functions as a complement of the matrix's verb and may be preceded (25) or not (26) by an infinitival complementizer (Cinf).

In Wolof linguistics, the matrix verb in constructions like (25-26) (e.g. *bëgg*) are called *verb operators* (Church, 1981; Voisin-Nouguier, 2006) and are classified into three types: (i) *primary*, (ii) *semi-primary operators*, and (iii) *secondary operators*. *Primary operators* are two-arguments verbs whose complement function can only be fulfilled by an infinitive phrasal verb complement, since they no longer exist as lexical verbs (Voisin-Nouguier, 2006). An example of a primary verb operator is *soog* 'start', as illustrated in (27).

- (27) a. *Awa soog a nelaw*  
 Awa start Cinf sleep  
 "Awa started to sleep."

*Semi-primary operators*, in contrast, are ambivalent verbs which can subcategorize for infinitival complements (25-26), but also for objects (28a) and non-finite complements (28b). *Secondary operators* — not discussed here — constitute a wide range of heterogeneous verbs which do not all exhibit the same syntactic properties.

- (28) a. *Faatu bëgg na [jëkkër]*  
 Faatu want +F.3sg husband  
 "Faatu wants a husband."  
 b. *Awa bëgg [mu jël xale bi]*  
 Awa want 3sg.subj take child the  
 "Awa wants him to take the child."

As we will see in section 4.3, the primary and semi-primary operators illustrate different cases of control relations (Mohanani, 1983; Kroeger, 2004). Typically, a control relation involves two things: (i) a particular argument (e.g. SUBJ, OBJ) of a subordinate clause is omitted; (ii) that missing argument is interpreted as referring to a particular argument (e.g. SUBJ, OBJ) of the main clause. For instance, in examples (25-26), the respective infinitive clause can be analyzed as containing an invisible subject (i.e. a *controllee*), which is identified with the overt subject of the main clause (i.e. the *controller Móodu*). The matrix verb (e.g. *bëgg*) illustrates an instance of a control verb, i.e. a superordinate verb that "controls" the arguments of a subordinate, infinitive verb (i.e. *lekk*).

## 4 Wolof morphosyntax within LFG

In the following, I present the most important grammatical structures of the constructions discussed in section 3 and point out how they are dealt with in the Wolof grammar. The section starts with the analysis of the determiner phrase (*DP*).

### 4.1 Word order

In the proposed c-structure analysis for Wolof, a *DP* is headed by *D*, which can be occupied by different kinds of determiners, including definite (6-7) and indefinite articles (8), interrogative determiners. For instance, the c-structure and f-structure of the *DP xale bi* "the child" in (6) are given in Figure 3.

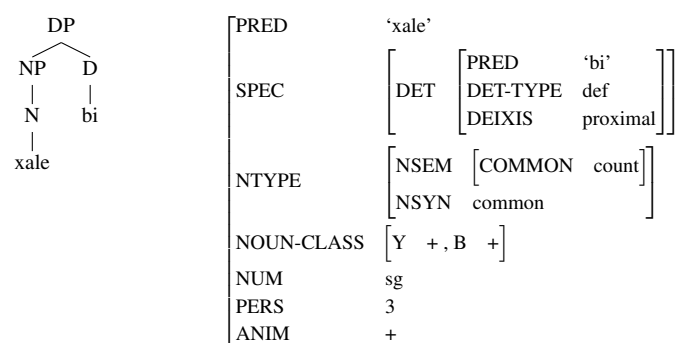


Figure 3: C-structure and f-structure of example (6)

The determiner introduces a *DET* feature under *SPEC* that indicates the semantic predicate ('bi'), the deixis (proximal) and the type of the determiner (e.g. definite). It also specifies the person and number of the structure. The *NP* is the f-structure co-head of *D* and contributes different syntactic and semantic features within the *DP*, including the basic syntactic and semantic type of the noun. Thus, the value of the attribute *NSYN* refers to a common noun in contrast to proper nouns and pronouns. Likewise, the semantic feature *NSEM* signals that the noun is a count noun which contrasts to mass nouns. Also, agreement between the determiner and the noun is controlled via a constraining equation — not displayed here — which, for instance, makes sure that the determiner *bi* agrees with the noun *xale* in the *B* class, i.e. a noun with the f-structure [B +].<sup>7</sup>

Furthermore, Wolof displays typologically mixed endocentric and exocentric organization (Bresnan, 2001). In endocentric constructions, the c-structure organization obeys the basic principles of X-bar (or  $\bar{X}$ ) (Jackendoff, 1977; Chomsky, 1986): a phrase is internally headed; a phrase and its head have the same category, but a different bar level. Following this model, a Wolof sentence is analyzed as an *IP* or *IPfoc*, depending on whether it contains focus information.<sup>8</sup> Such a clause is assumed to be headed by a verbal function category *I<sup>o</sup>* or *Ifoc* (for Inflection).

The Wolof LFG also assumes the exocentric category *S* for non-focused clauses. The *S* category does not obey the constraints of  $\bar{X}$  theory. It behaves as a maximal phrase, but it has no c-structure head, and it can dominate phrases of any category or bar level.

The analysis of a non-focused *IP* is illustrated with the optative clause in (29). Figure 4 shows the c-structure and simplified f-structure associated with (29).

Likewise, it is motivated to assume the exocentric *S* (or more precisely *Sfoc*) category for focused constructions. For instance, let us consider the non-subject focus clause in (30), in which morphological information rather than phrase structure position determines the grammatical function of a phrase. As the c-structure in Figure 5 shows, the finite auxiliary (i.e. *la*) occurs in *Ifoc* whose complement is analyzed as *Sfoc* (not *VPfoc*), since the node dominates the *DP* subject (i.e. *xale yi*) as well as *VPfoc*. This illustrates again a typical case of an exocentric phrase. Also, the specifier of *IPfoc* simultaneously bears the function subject and the discourse focus, as the indexes in the f-structure indicate.

<sup>7</sup>See Dione (2014) for more details on the f-structure analysis of Wolof noun classes.

<sup>8</sup>*IPfoc* is singled out as a formal category for purely parsing efficiency reasons. Theoretically, focus is a functional feature and is equally well handled in the f-structure.

- (29) *Na xale y-i lekk jën wi*  
 +Opt.3 child cl.pl-DFP eat fish the.sg  
 “May the children eat the fish.”

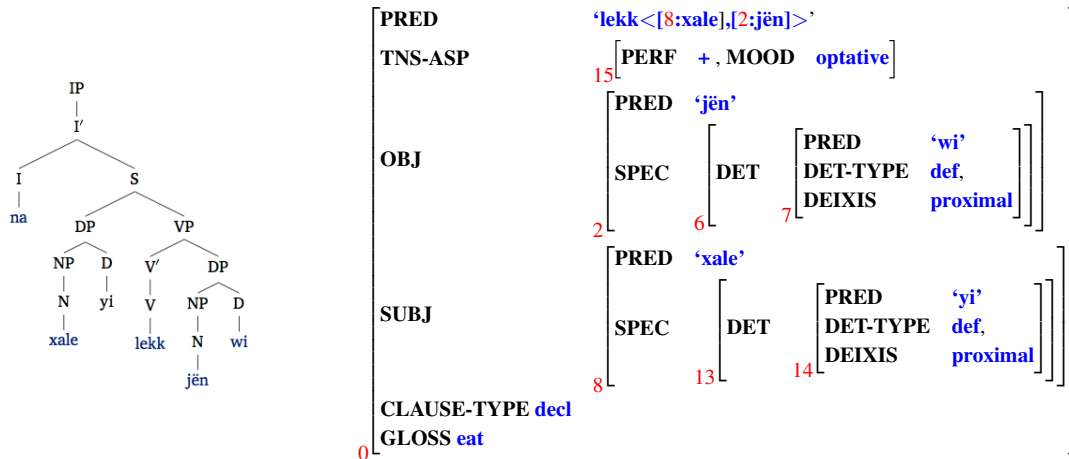


Figure 4: C-structure and simplified f-structure of optative clause (29)

- (30) *Jën wi la xale yi lekk*  
 Fish the.sg NSFOC.3 child cl.pl-DFP eat  
 “It’s the fish that the children have eaten.”

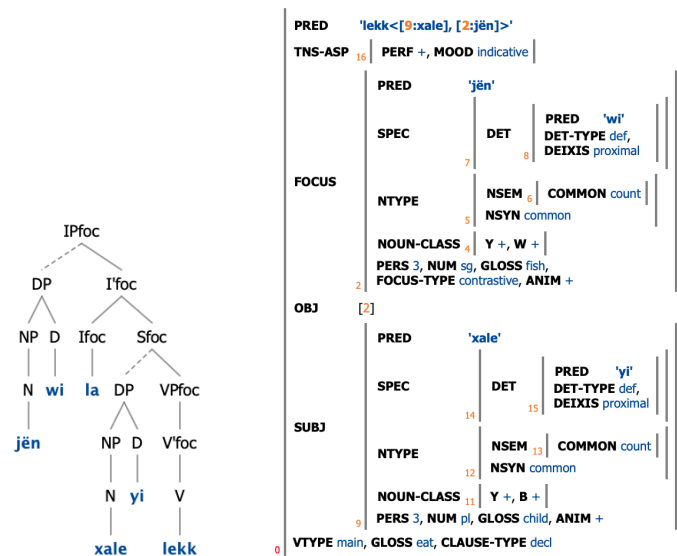


Figure 5: C- and f-structure of the non-subject focus clause (30)

Narrative clauses (22) are analyzed as *S* (see Figure 6), as the verb in such clauses is unspecified for finiteness, person inflection and polarity.

Besides the *I* and *D* categories, Wolof makes use of the functional category *C* (for complementizers). Accordingly, *IP* may be dominated by *CP*, i.e. a sentence with a complementizer or a displaced phrase in sentence-initial position. Subordinate phrases with initial or without overt complementizer are *CP*s, as in (31), which is headed by *ne* ‘that’.

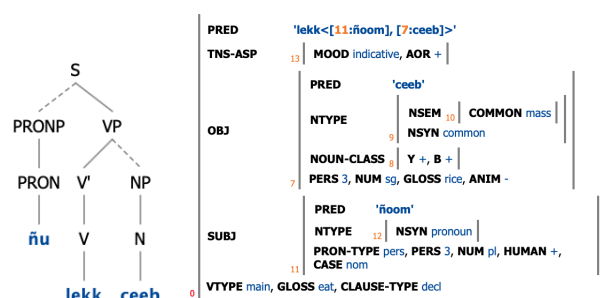


Figure 6: C- and f-structure of the narrative clause (22)

- (31) *Xam naa ne Móodu dem na*  
 know +F.1sg COMP Móodu go +F.3sg  
 “I know that Móodu has left.”

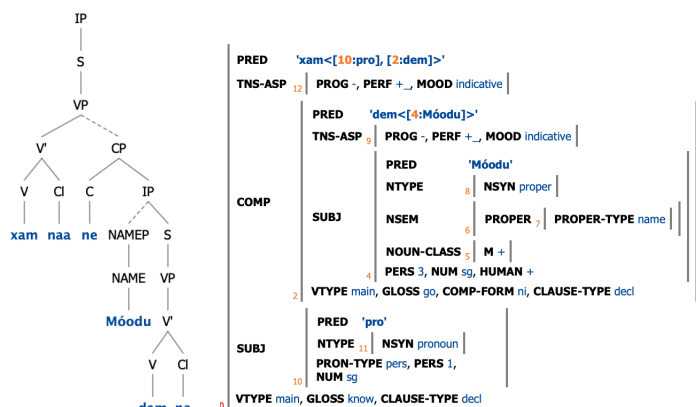


Figure 7: C- and f-structure of example (31)

## 4.2 Pro-drop

In the LFG architecture, pronoun incorporation can be analyzed as follows: an incorporated pronoun or pronominal inflection is treated as “a bound morpheme that specifies a complete pronominal f-structure. The functional specification of a pronoun is incorporated with the functional specifications of the stem to which the morpheme is bound” (Bresnan, 2001, pages 145-146). The functional specifications of a pronoun may include various features like semantic features, binding features, case and agreement features.

Under the pro-drop analysis proposed by Bresnan and Mchombo (1987), the subject marker (SM) specifies subject agreement features to the functional analysis of the clause. As Figure 8 shows, the subject marker optionally provides a PRED ‘pro’. In case there is no overt NP subject, this referential option is realized. Otherwise, if there is an overt NP subject, the PRED feature will be supplied by the full NP. This is in order to prevent feature clash as PRED features in LFG are not subject to unification (Butt, 2007). The *agr* feature is a metavariable over features like person, number and gender of the subject.

The Wolof grammar follows the standard LFG approach to pro-drop. As (33) shows, the verb specifies a great deal of information about its subject argument, including an optional PRED value, person and number. Clausal negation is specified by the morpheme *-u*. This pro-drop analysis allows the f-structure for a sentence like (32) to satisfy Completeness (Kaplan and

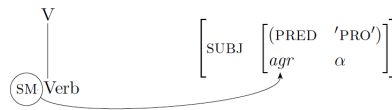


Figure 8: Classic LFG analysis of pro-drop (Bresnan and Mchombo, 1987)

Bresnan, 1982), as the thematic SUBJ required by the verb is present and has a PRED. Completeness means that all the governable grammatical functions required by the PRED of the f-structure should have a value in the f-structure. As Figure 9 shows, a Wolof sentence may consist simply of a verb, with no overt subject phrase present at c-structure.

- (32) *Dem-u-ñu.*  
 go-NEG-3pl  
 “They haven’t left.”
- (33) *dem-u-ñu V* (↑ PRED)=‘dem<(↑ SUBJ)>’  
 ((↑ SUBJ PRED)=‘pro’)  
 (↑ SUBJ PERS)=3  
 (↑ SUBJ NUM)=pl  
 (↑ NEG)=+.

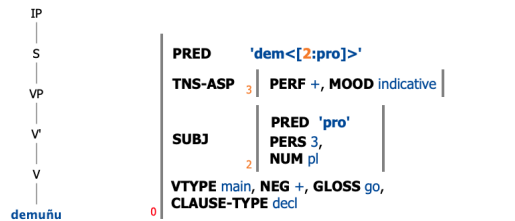


Figure 9: C- and f-structure analysis for the sentence (32).

In contrast to languages like Wolof and Chicheŵa, there are other languages that allow arguments to be dropped without any corresponding agreement morphology. In ParGram, there are at least two of such languages: Japanese (Masuichi et al., 2003) and Urdu (Butt, 2007), which propose a different approach to pro-drop.

Like many South Asian languages, Hindi/Urdu has the ability to pro-drop any and all arguments (Butt, 2007). However, in these languages pro-drop does not necessarily correlate with agreement and pronominal incorporation (Butt, 2007), but rather with discourse. According to Butt and King (1997), only old information such as continuing topics or background information is dropped. To capture this fact, an i(nformation) structure is assumed along with the usual c- and f-structure. The i-structure projects from the c-structure, allowing for a clean separation of predicate-argument information from information structure. The LFG analysis proposed for pro-drop in Hindi/Urdu is illustrated by the Hindi sentences given in (34a-34b) and their respective associated f-structures in Figures 10 and 11. The sentence (34b) has a pro-dropped subject, which is a continuing topic from the previous utterance, as indicated by coreference in the i-structure in Figure 12.

- (34) a. *mē̃ bais bars=se yahā rah rah-a hū*  
 I.Nom twenty-two year=from here live Prog-M.Sg be.Pres.1.Sg  
 ‘I<sub>topic</sub> have been living here for 22 years.’
- b. *rozana is hi sarak=se gūzar-ta hū*  
 daily this Emph street.F=from pass-Impf.M.Sg be.Pres.1.Sg  
 ‘Daily (I<sub>cont.topic</sub>) go through this street.’

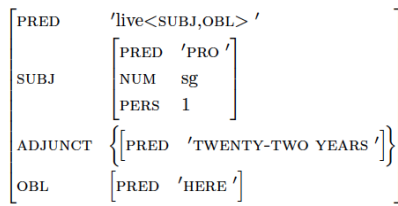


Figure 10: F-structure for (34a)

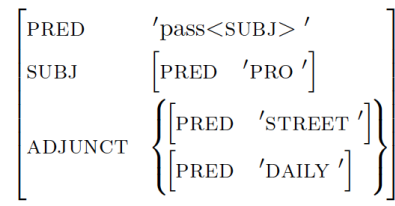


Figure 11: F-structure for (34b)

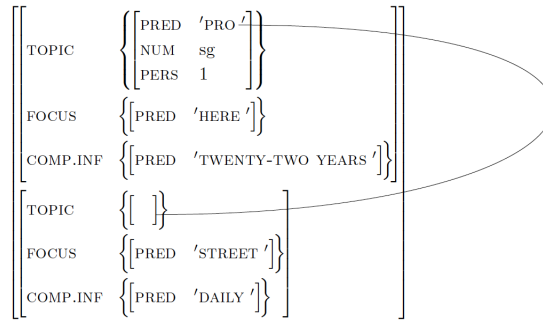


Figure 12: Information structure for (34a) and (34b)

The pro-drop analysis proposed for Japanese makes use of Optimality Theory (OT) constraints (Frank et al., 1998; Frank et al., 2001). Based on the literature on Optimality Theory (Prince and Smolensky, 1993), Frank et al. (2001) proposed a new projection referred to as *o*-structure as an extension of the LFG projection architecture. The *o*-structure is used to determine a preference ranking on the set of analyses for a given input sentence. A relative ranking is specified for the constraints that appear in the *o*-projection. This ranking serves to determine the winner among the competing candidates.

To illustrate how the OT mechanism is used in the Japanese LFG grammar (Masuichi et al., 2003) to handle pro-drop, let us consider the lexical annotations of the verb *yomu* ‘read’ in (35). This example first states that the verb has a semantic predicate *yomu* which subcategorizes for a subject and an object. Then, the annotations in the next two rows specify the predicate and pronoun type of the subject be ‘pro’ and null, respectively. Likewise, (↑ OBJ PRED)=‘pro’ and (↑ OBJ PRON-TYPE)=null refer to the predicate and pronoun type of the object. This means that both arguments (subject and object) can be dropped. However, these annotations work only if no constituent that can be subcategorized for by the verb *yomu* exists in the input. This is achieved by using the OT mark “ProDrop:OT” to indicate this preference. In order for this preference rule to work as desired, the OT constraint is set at the lowest level in the Japanese grammar.

- (35) *yomu*      V    (↑ PRED)=‘yomu<(↑ SUBJ)(↑ OBJ)>’  
 (↑ SUBJ PRED)=‘pro’  
 (↑ SUBJ PRON-TYPE)=null  
 (↑ OBJ PRED)=‘pro’  
 (↑ OBJ PRON-TYPE)=null  
 ProDrop: OT.



### 4.3 Functional and anaphoric control

Wolof has several types of control constructions, i.e. constructions “in which either syntactic or lexical constraints require coreference between an argument of the matrix clause and an argument of a subordinate or modifying adjunct clause” (Dalrymple, 2001, page 313). Wolof control constructions are related to the non-finite verbal complements and infinitival clauses discussed in section 3.2.

#### 4.3.1 Raising verbs in Wolof

As noted above, Wolof has verbs referred to as *primary operators* (Church, 1981) that no longer exist as lexical verbs. Such verbs typically subcategorize for two argument functions: a subject and a complement which can only be an infinitival one (Voisin-Nouguier, 2006). This presents a strong evidence that *primary operators* exemplify a typical case of obligatory control relation.

Nevertheless, the status of the subject of such verbs, e.g. *Awa* in (36-37), should be examined more closely. On the one hand, it is not a semantic argument of the verb, e.g. *soog*. As the contrasts between (36)<sup>9</sup> and (37)<sup>10</sup> vs. (38)<sup>11</sup> show, the verb of the subordinate clause (e.g. *nelaw* vs. *taw*) determines the semantic property of the subject (e.g. *Awa*), but not the verb of the matrix clause (e.g. *soog*). In other words, the latter verb imposes no semantic constraints on the SUBJ as a “raised” argument. Only the infinitival complement (e.g. *nelaw*) is semantically related to the main verb. On the other hand, this SUBJ is a syntactic argument of the matrix verb (*soog*), as well as of the subordinate verb (*nelaw*). Therefore, it must be part of the argument structure and PRED feature value of the matrix verb.

(36) *Awa soog a nelaw*  
 Awa start Cinf sleep  
 “So, Awa starts to sleep.”

(37) #*Awa soog a taw*  
 Awa start Cinf rain  
 “So, Awa starts to rain.”

(38) *Mu soog a taw*  
 PRONexpl start Cinf rain  
 “So, it starts to rain.”

There are similarities between raising verbs in English and the Wolof verbs classified as *primary operators*. For instance, example (36) is similar to the English sentence (39), in which there are two verbs, but there is only one thematic role involved. The subject is an argument in the subordinate clause (*yawn*), but not an argument in the matrix clause (*seem*) (Dalrymple, 2001; Falk, 2001). The complement of *seem* is a functionally controlled open complement (XCOMP) function (Bresnan, 1982; Butt et al., 1999a), i.e. a complement whose subject is obligatorily functionally controlled from outside the clause. In functional control, the subject of the matrix verb is identical to the subject of the subordinate verb, i.e. syntactic restrictions imposed by the subordinate verb must be maintained when the subject is raised (Dalrymple, 2001). This identity is resolved by a functional control equation on the lexical entry of the raising verb, as shown in Figure 13.

<sup>9</sup>Example (27) is repeated in (36).

<sup>10</sup>As indicated by the # symbol, (37) is semantically ill-formed (but not ungrammatical).

<sup>11</sup>Note that the form *mu* is ambiguous between an expletive (38) and a personal pronoun (41b).

(39) David seemed to yawn.

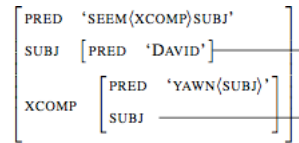


Figure 13: F-structure of (39)

Besides these similarities, we may need to consider another source of evidence for assuming that the subject of a Wolof *primary operator* like *soog* is not a semantic argument of that verb. In fact, most of the Wolof primary operators are better translated into English as adverbs or adverbial phrases. Specific examples include e.g. *faral* “(be) used to”, *mas* “(have) once”, *soog* “(to do) for the first time”, *waaj* “(be) about to”, *xaw* “almost (do)”. As can be observed from the translations, these verbs show similarity with adverbs in that they provide a description of how, when, in what manner and to what extent something is done or happens.

Based on these observations, I assume that this class of Wolof verbs can be analyzed as raising verbs. This means that these predicates do not assign any semantic role to their subjects, and place no semantic restrictions on their subjects, except the restrictions required by the embedded verb. Moreover, Wolof primary operators exemplify functional control. Following this assumption, the subject of the primary verb operator functionally controls the SUBJ of the embedded clause. In turn, the embedded clause bears the XCOMP function of the matrix clause. This means that the SUBJ of the verb *soog* ‘start to do’ in (37) is required to be the same f-structure as the SUBJ of the subordinate XCOMP, as the indexes in the f-structure in Figure 14 indicate. Also, as shown in Figure 14, the subject of the main predicate is assumed to be a non-semantic argument of that predicate. Therefore, it appears outside the angled brackets.

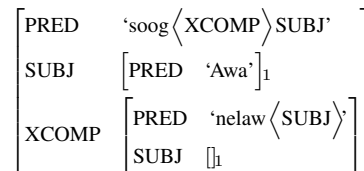


Figure 14: Simplified f-structure analysis of example (36)

#### 4.3.2 Equi verbs in Wolof

In contrast to primary operator verbs, a semi-primary operator verb like *bëgg* in (40) can subcategorize for infinitival (40) or non-finite complements (23a) as well as for nominal objects (28a).

(40) *Awa bëgg na lekk*  
 Awa want +F.3sg eat  
 “Awa wants to eat.”

Example (40) is similar to *equi* sentences in English, such as *David tried to leave*, which can be interpreted as “meaning that David tried to bring about a situation where David leaves” (Dalrymple, 2001, page 324). As in English, in Wolof, the arguments of the semi-primary operator

verbs are syntactic and semantic arguments (unlike the situation with raising verbs). Thus, unlike primary operators, all arguments of the semi-primary operators are semantic arguments. The subject of semi-primary operator verbs has two thematic roles. It is a thematic argument of the main verb and also a thematic argument of the complement. These observations suggest that it is legitimate to assume semi-primary operators in Wolof to be *equi* verbs.

According to Dalrymple (2001), English *equi* verbs exemplify anaphoric control, while English raising verbs exhibit functional control. For Wolof, however, I assume that *equi* verbs can participate in either functional or anaphoric control. In this context, the difference between functional and anaphoric control should briefly be pointed out. In an anaphoric control construction, syntactic restrictions are not imposed on the subject of the subordinate complement. Although the anaphorically controlled subject of the subordinate complement and the matrix controller are semantically related by an anaphoric binding relation, the former is syntactically independent from the latter.

As explained by Dalrymple (2001, page 329), some languages like Tagalog “have two types of *equi* verbs, some specifying anaphoric control and some specifying functional control”. Similarly, I assume that Wolof has indeed two different types of *equi* constructions and therefore falls into this category of languages. The first type of *equi* constructions involves functional control of a subject argument in the complement clause, as illustrated in (40). The corresponding f-structure analysis is shown in Figure 15. Unlike raising verbs, the controller in an *equi* construction is semantically as well as syntactically selected by the verb. Notationally, this is reflected in the fact that the SUBJ of the *equi* verb *bëgg* in (40) appears inside rather than outside the angled brackets in the semantic form.

$$\left[ \begin{array}{l} \text{PRED} \quad \text{'bëgg} \langle \text{SUBJ, XCOMP} \rangle \\ \text{SUBJ} \quad \left[ \text{PRED} \quad \text{'Awa'} \right]_1 \\ \text{XCOMP} \quad \left[ \begin{array}{l} \text{PRED} \quad \text{'lekk} \langle \text{SUBJ} \rangle \\ \text{SUBJ} \quad \left[ \right]_1 \end{array} \right] \end{array} \right]$$

Figure 15: Simplified f-structure analysis of example (40)

The second type of *equi* constructions are found, for instance, in object control constructions like (41).

- (41) a. *Xale yi aaye na-ñu Móodu (mu) dem*  
 child the.pl prevent +F-3pl Móodu 3sg.subj leave  
 “The children prevented Móodu from leaving.”
- b. *Xale yi aaye na-ñu ko (mu) dem*  
 child the.pl prevent +F-3pl 3sg.obj 3sg.subj leave  
 “The children prevented him from leaving.”

Examples (41a-41b) involve anaphoric control of the non-subject argument which is the subject argument of the complement clause. In (41a), the controller of the SUBJ of *dem* ‘leave’ is the OBJ of the matrix verb, *Móodu*. The simplified f-structure analysis corresponding to example (41a) is shown in Figure 16. Similarly, in (41b), the matrix object (i.e. *ko*) anaphorically

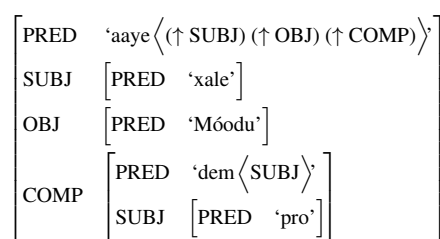


Figure 16: Simplified f-structure analysis of example (41a)

controls the SUBJ argument in the subordinate clause (i.e. *mu*). This exemplifies cases of non-coreference between the subject of the matrix (i.e. *xale yi*) and the subject of the subordinate clause (i.e. *mu*). In these examples of object control constructions, the subject marker (e.g. *mu*) is optional. However, the object in the matrix clause, which surfaces as a clitic *ko* in (41b) or a lexical object *Móodu* in (41a), is never optional, as expected with null objects (Torrence, 2005).<sup>12</sup>

Thus, the Wolof case serves as support for the statement that “equi verbs involving functional as well as anaphoric control can be found, even within the same language” (Dalrymple, 2001, page 330).

### 4.3.3 Anaphoric control in infinitival clauses

In Wolof, infinitival clauses like the bracketed phrases in (42), occur relatively frequently.

- (42) a. [Wat]            a      gën      [wef]  
 shave.the.hair SFOC surpass pull.out.the.hair  
 “It is better to shave the hair than pulling it out.” (Prov.)  
 (of two evils choose the lesser)
- b. [Taxawu yokkute réew m-i]      yomb-ul  
 assist      progress country cl-DFP be.easy-NEG.3sg  
 Lit.: “Assisting the progress of the country is not easy.”  
 “It is not easy to help the country develop.”

Constructions in (42) are similar to the English examples (43-44) in that they “have a nonovert subject which is not functionally controlled” (Butt et al., 1999a, page 53). However, unlike in English, which requires the clause to be headed by *to* (43) or be a gerundive *VP* (44), the clauses in (42) are bare infinitives. They can express core argument functions, including SUBJ (42a-42b), OBJ (42a), complement of a predicate, etc.

- (43) [To clone dinosaurs] would please the geneticist. (Falk, 2001, page 121)
- (44) [Pinching those elephants] was foolish. (Butt et al., 1999a, page 54)

In (42), the head of the constituents in brackets is a verbal base. The verb can be intransitive (42a) and transitive (42b). As a transitive verb, it takes its canonical arguments SUBJ and OBJ, while SUBJ need not be expressed. In (42b), for instance, the verb *taxawu* takes its canonical object argument. In the current Wolof grammar, these clauses are analyzed as infinitival VPs,

<sup>12</sup>In (41b), *mu* is subject marker rather than object, since the corresponding object form is *ko*.

*VPinf*. As such, the verb and its argument — e.g. “[*Taxawu yokkute réew mi*]” in (42b) — form a single constituent that bears the subject function of the clause. Similarly, the infinitival *VP* [*wef*] in (42a) is assumed to bear the non-subject argument of the main verb *gën* ‘surpass’ in (42a).

Note that the infinitive verbs (e.g. *taxawu*, *wat*, *wef*) of the subordinated clauses in (42) do not form a complex predicate with the matrix verb (Voisin-Nouguier, 2006). The infinitive clauses bear a subject or an object function, rather than an (open or closed) complement function for the verb *gën*. The finite verbs and the infinitive verbs refer to completely distinct events, which are not influencing each other. In (42), there is, for instance, no control of the main verb *yomb* or *gën* imposed on any arguments of the infinitive verbs *taxawu*, *wat*, *wef*, etc. The non-subject argument in the construction in (42a), i.e. *wef*, is treated as an OBJ, rather than COMP, of the verb. The motivation behind this analysis is that the verb *gën* typically subcategorizes for object arguments, but not for sentential complements.

For Wolof, I assume these clausal arguments to be anaphorically controlled: the unexpressed subject in these constructions refers to an arbitrary entity (i.e. a null referent) which must be determined from the context of the utterance. Accordingly, clauses like (42a) are analyzed as shown in Figure 17. The f-structure analysis encodes this relation in terms of an anaphorically controlled NULL\_PRO with PRON-TYPE null whose referent is yet to be determined.

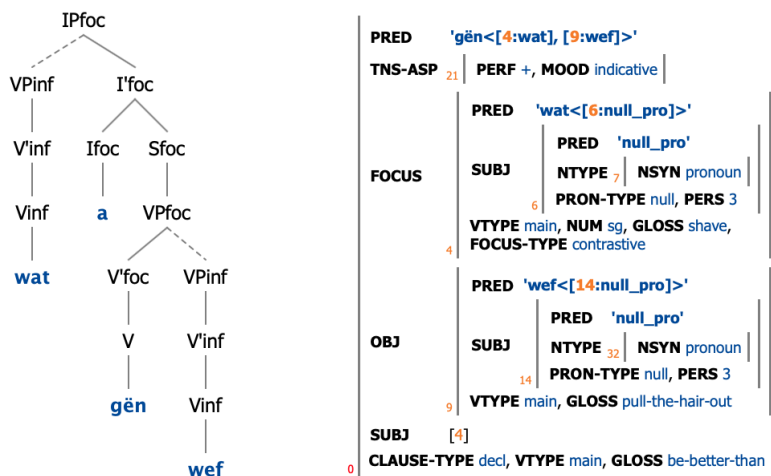


Figure 17: Anaphoric control in Wolof infinitive clauses with a null referent subject

The problem of clausal complements in general and what grammatical function they should take is a long-standing debate in LFG. Alsina et al. (2005) suggested to analyze clausal complements CPs as OBJ/SUBJ, thereby removing COMP from the inventory of the grammatical functions assumed in LFG. This was in order to eliminate redundancy, thus simplifying the LFG framework. In contrast, Dalrymple and Lødrup (2000) pointed out that an adequate account of so called mixed languages (e.g. English, German, and Swedish) requires a distinction between clausal complements that are objects and those that are non-objects. To investigate such a distinction, Lødrup (2004) defined three main criteria: alternation with a DP object, unbounded dependencies and passivization. First, it is necessary for an object clausal complement to be able to alternate with a (thematic) DP object. A clausal complement that does not alternate with a DP object is therefore considered as a non-object. Second, a clausal complement that cannot

topicalize is not an object, but this does not mean that only object clausal complements can topicalize. Third, an object corresponds to a subject in the passive, while a non-object clausal complement does not. Using examples from Norwegian, Lødrup (2004, pages 70-71) showed that infinitival complements in that language are objects in the majority of cases. For instance, the complement clause (i.e. the bracketed constituent in (45)) of the verb *akseptere* ‘accept’ alternates with a DP object (46), topicalizes (47) and corresponds to a subject in the passive (48).

(45) *De har akseptert [å betale høyere skatt].*  
they have accepted to pay higher tax

(46) *De har akseptert dette.*  
they have accepted this

(47) *[Å betale høyere skatt] har de akseptert.*  
to pay higher tax have they accepted

(48) *[Å betale høyere skatt] er blitt akseptert.*  
to pay higher tax has been accepted

As in Norwegian, the Wolof infinitival complements discussed in this section are mostly objects or subjects. To verify this, we can apply similar tests for Wolof using these criteria. For the second criterion, we can use both topicalization and clefting as instances of unbounded dependencies. For the third criterion, we can use a construction with a 3PL impersonal pronoun in subject position, as this is the way to expressive passive in Wolof. Thus, sentence (49) can also have a passive interpretation. As (49-52) show, the infinitival complement, i.e. the bracketed constituent in (49), can alternate with a DP object (50), topicalize (51) and can be put into focus through clefting. In topicalized structures like (51), the obligatory resumptive pronoun (e.g. *ko*) corefers with the topicalized infinitival complement.

(49) *Bëgg na-ñu [taxawu yokkute réew m-i].*  
want +F-3pl assist progress country cl-DFP  
‘They want to help the country develop’

(50) *Bëgg na-ñu ko.*  
want +F-3pl 3sg.Obj  
‘They want this’

(51) *[Taxawu yokkute réew m-i] bëgg na-ñu ko.*  
assist progress country cl-DFP want +F-3PL 3sg.Obj  
‘Help the country develop, they want it’

(52) *[Taxawu yokkute réew m-i] la-ñu bëgg.*  
assist progress country cl-DFP NSFOC-3pl want  
‘Help the country develop is what they want’

The main difference between the Norwegian and Wolof infinitival complements is that the former are headed by an infinitival complementizer ‘å’, whereas the latter have bare infinitive form. Accordingly, the Norwegian constituents are typically analyzed as *CPs* at c-structure, while the Wolof ones are treated as a special category of *VP* (i.e. *VPinf*).

## 5 Conclusion

This paper has presented the construction of a computational grammar for Wolof using the LFG model. First, I provided a general overview of the LFG theory and parsing with XLE. I have discussed a range of constructions related to the clausal/verbal system of Wolof and shown how these are handled in an LFG setting. In the discussion of pro-drop constructions, LFG's traditional analysis with an empty grammatical function with the PRED value 'PRO' is chosen for the grammatical representation of no overt subjects. This solution retains the usual subcategorization requirements without introducing unnecessary empty nodes into the phrase structure. This classic approach to pro-drop is contrasted with analyses proposed for other languages such as Urdu and Japanese in which arguments can be dropped without any corresponding agreement morphology.

Furthermore, various types of control constructions in Wolof are examined. Wolof primary operators are analyzed as raising verbs which exemplify functional control, while semi-primary operators are assumed to be *equi* verbs that can participate in either functional or anaphoric control. This serves as support for the assumption that *equi* verbs involving functional as well as anaphoric control can be found within the same language.

## References

- Alex Alsina, Tara Mohanan, and KP Mohanan (2005).**  
“How to Get Rid of the COMP”. In: *The Proceedings of the LFG '05 Conference*. Ed. by Miriam Butt and Tracy Holloway King. University of Bergen, Norway.
- Matthew Baerman, Dunstan Brown, and Greville G Corbett (2005).**  
*The Syntax-Morphology Interface: A Study of Syncretism*. Cambridge University Press.
- Kenneth R. Beesley and Lauri Karttunen (2003).**  
*Finite State Morphology*. Stanford, CA: Center for the Study of Language and Information.
- Joan Bresnan (1982).**  
“Control and Complementation”. In: *The Mental Representation of Grammatical Relations*. Ed. by Joan Bresnan. Cambridge, MA: The MIT Press, pp. 282–390.
- Joan Bresnan (2001).**  
*Lexical-Functional Syntax*. Oxford: Blackwell.
- Joan Bresnan and Sam A. Mchombo (1987).**  
“Topic, pronoun, and agreement in Chicheŵa”. In: *Language* 63.4, pp. 741–782.
- Miriam Butt (2007).**  
“The role of pronominal suffixes in Punjabi”. In: *Architecture, Rules, and Preferences: Variations on Themes by Joan W. Bresnan*. Ed. by Annie Zaenen et al. Stanford, CA: CSLI Publications, pp. 341–368.
- Miriam Butt and Tracy Holloway King (1997).**  
“Null elements in discourse structure”. In: *Seminar on NULLs, Delhi University*. Ed. by K.V. Subbarao. To appear. Motilal Banarasidass.
- Miriam Butt et al. (1999a).**  
*A Grammar Writer's Cookbook*. Stanford, CA: CSLI.
- Miriam Butt et al. (1999b).**  
“Writing Large-Scale Parallel Grammars for English, French, and German”. In: *Proceedings of the LFG99 Conference*. Stanford, CA: CSLI Publications, pp. 1–59.
- Miriam Butt et al. (2002).**  
“The Parallel Grammar Project”. In: *Proceedings of the COLING02 Workshop on Grammar Engineering and Evaluation*. Vol. 15. Association for Computational Linguistics, pp. 1–7.
- Noam Chomsky (1986).**  
*Barriers*. The MIT Press.
- Eric D. Church (1981).**  
*Le système verbal du wolof*. Faculté des Lettres et Sciences Humaines, Université de Dakar.
- Dick Crouch et al. (2019).**  
XLE Documentation. On-line documentation, Palo Alto Research Center (PARC).
- Mary Dalrymple (2001).**  
*Lexical-Functional Grammar (Syntax and Semantics, Volume 34) (Syntax and Semantics)*. Academic Press.
- Mary Dalrymple and Helge Lødrup (2000).**  
“The grammatical functions of complement clauses”. In: *Proceedings of the LFG00 Conference*. Ed. by Miriam Butt and Tracy Holloway King. CSLI Publications.
- Cheikh Bamba Dione (2012).**  
“A Morphological Analyzer For Wolof Using Finite-State Techniques”. In: *Proceedings of the 8th LREC*. ELRA.



**Cheikh Bamba Dione (2013).**

“Handling Wolof Clitics in LFG”. In: *Challenging Clitics*. Ed. by Christine Meklenborg Salvesen and Hans Petter Helland. Amsterdam: John Benjamins Publishing Company.

**Cheikh Bamba Dione (2014).**

“LFG parse disambiguation for Wolof”. In: *Journal of Language Modelling* 2.1, pp. 105–165.

**Cheikh Bamba Dione (2017).**

“Finite-State Tokenization for a Deep Wolof LFG Grammar”. In: *Bergen Language and Linguistics Studies* 8.1.

**Yehuda Falk (2001).**

*Lexical Functional Grammar: an Introduction to Parallel Constraint-Based Syntax*. Stanford, CA: CSLI Publications.

**Anette Frank et al. (1998).**

“Optimality Theory Style Constraint Ranking in Large-scale LFG Grammars”. In: *The Proceedings of the LFG98 Conference*. Ed. by Miriam Butt and Tracy Holloway King. Stanford, CA: CSLI Publications.

**Anette Frank et al. (2001).**

“Optimality Theory Style Constraint Ranking in Large-scale LFG Grammars”. In: *Formal and Empirical Issues in Optimality Theoretic Syntax*. Ed. by Peter Sells. Stanford, CA: CSLI Publications.

**Ray Jackendoff (1977).**

“X-bar Syntax: A Study of Phrase Structure”. In: *Linguistic Inquiry Monograph Two*.

**Omar Ka (1994).**

*Wolof Phonology and Morphology*. Lanham, Maryland: University Press of America.

**Ron Kaplan and Joan Bresnan (1982).**

“Lexical-Functional Grammar: A Formal System for Grammatical Representation”. In: *The Mental Representation of Grammatical Relations*. Ed. by Joan Bresnan. Cambridge, MA: MIT Press, pp. 173–281.

**Ronald M. Kaplan et al. (2004).**

“Integrating Finite-State Technology with Deep LFG Grammars”. In: *Proceedings of the ESSLLI’04 Workshop on Combining Shallow and Deep Processing for NLP*.

**Paul Kroeger (2004).**

*Analyzing syntax: a lexical-functional approach*. Cambridge University Press.

**Helge Lødrup (2004).**

“Clausal complementation in Norwegian”. In: *Nordic Journal of Linguistics* 27.1, pp. 61–95.

**Hiroshi Masuichi et al. (2003).**

“Japanese parser on the basis of the Lexical-Functional Grammar formalism and its evaluation”. In: *Proceedings of PACLIC17*, pp. 298–309.

**Fiona McLaughlin (1997).**

“Noun classification in Wolof: When affixes are not renewed”. In: *Studies in African Linguistics* 26.1.

**Karuvannur P Mohanan (1983).**

“Functional and anaphoric control”. In: *Linguistic Inquiry* 14.4, pp. 641–674.

**Alan Prince and Paul Smolensky (1993).**

*Optimality Theory: Constraint interaction in Generative Grammar*. Tech. rep. Rutgers University Center for Cognitive Science.

**Stéphane Robert (1991).**

“Approche énonciative du système verbal: le cas du wolof”. In: *Editions du CNRS*.

**Stéphane Robert (1996).**

“Aspect zéro et dépendance situationnelle: l'exemple du wolof”. In: *Dépendance et intégration syntaxique (subordination, coordination, connexion)*, pp. 153–161.

**Sebastian Sulger et al. (2013).**

“ParGramBank: The ParGram Parallel Treebank”. In: *Proceedings of the 51th Annual Meeting of the Association for Computational Linguistics (ACL 2013)*. Sofia, Bulgaria, pp. 759–767.

**Khady Tamba, Harold Torrence, and Malte Zimmermann (2012).**

“Wolof quantifiers”. In: *Handbook of Quantifiers in Natural Language*. Springer, pp. 891–939.

**Harold Torrence (2003).**

“Verb Movement in Wolof”. In: *UCLA Working Papers in Linguistics*, pp. 1–31.

**William Harold Torrence (2005).**

“On the Distribution of Complementizers in Wolof”. PhD thesis. Los Angeles, CA: University of California.

**Robert D Van Valin and Randy J La Polla (1997).**

*Syntax: Structure, Meaning, and Function*. Cambridge University Press.

**Sylvie Voisin-Nouguier (2006).**

“L'infinitif en wolof”. In: *Les formes non finies du verbe*. Ed. by C. Blanvillain O. & Guilmier. Presses Universitaires de Rennes, pp. 61–83.

**Anne Zribi-Hertz and Lamine Diagne (2002).**

“Clitic placement after syntax: evidence from Wolof person and locative markers”. In: *Natural Language & Linguistic Theory* 20.4, pp. 823–884.

**Anne Zribi-Hertz and Lamine Diagne (2003).**

“Défiance flexionnelle et temps topical en wolof”. In: *Typologie des langues d'Afrique et universaux de la grammaire, volume 2: benue-kwa, wolof*. Ed. by Patrick Sauzet et Anne Zribi-Hertz. L'Harmattan, pp. 205–231.