

**DEGEMA AND THE STRING INTERFACE**

Tina Bögel  
University of Konstanz

Proceedings of the LFG14 Conference

Miriam Butt and Tracy Holloway King (Editors)

2014

CSLI Publications

<http://csli-publications.stanford.edu/>

## Abstract

This paper reports on some of the phonological properties of Degema, a language spoken in the south of Nigeria, and focusses on the occurrence of endoclitics, which allow for interesting insights into the interplay of syntax, lexicon, postlexical phonological rules and prosody. This phenomenon presents a problem for current linguistic analysis, specifically to the concepts of modularity and lexical integrity, and requires a clear positioning of the different modules and the respective ‘strings’. On the basis of the Degema data, this paper proposes an architecture that allows for the analysis of endoclitisis, respects modularity and keeps the principle of lexical integrity intact.

## 1 Introduction

Recent years have seen several papers on the nature of the syntax-prosody interface within LFG (Bögel et al. 2009, Dalrymple and Mycock 2011, Bögel 2012, a.o.). The current general assumption is that the string is at the heart of the interface, incorporating (at least) two representations: the p-string which represents the phonological form of the string, and the s-string, which represents the syntactic side of the string. Both representations are interfaced by a multi-layered lexicon, which includes information on the phonological and syntactic properties of a lexical entry. However, the exact nature and position of the p-string is part of an ongoing debate, probably also owed to the fact that only little has been written about the (s-)string as such in the past. This paper aims to contribute to the discussion.

In general, the syntactically unparsed, but tokenized string is simply taken to be mapped to c-structure via the relation  $\pi$  (Kaplan 1987, Asudeh 2006), where c-structure represents the ‘linear order and hierarchical structure’ of that string (Dalrymple 2001, Asudeh et al. 2013). A slightly different view is taken by Asudeh (2009) who defines the (s-)string as a ‘representation of linear phonology’, thus assuming phonological and syntactic linear order to be parallelly represented within the string. Another recent analysis has adjusted the notion of linear order in that second position clitics were allowed to be moved in order to correctly analyse the associated syntactic information (Bögel et al. (2009), following Halpern (1995)), which implies a non-parallelity between the two sides of the string.

With reference to Degema, a language spoken in the South of Nigeria, this paper aims to look at these string concepts by analysing a notoriously difficult linguistic phenomenon: endoclitisis. Endoclitisis presents a challenge on several levels: First, the occurrence of a clitic element within the stem of another syntactic element is an obvious violation of the *Principle of Lexical Integrity*. Second, the analysis of endoclitics involves several aspects of the grammar: syntax, prosodic structure, postlexical phonological rules and the lexicon. It is thus also a discussion

---

<sup>†</sup>I would like to thank Dr. Ethelbert Emmanuel Kari, who patiently answered all my questions and provided the recorded data. I would also like to thank Miriam Butt, Tracy H. King, Frans Plank, the LFG audience in Ann Arbor, and an anonymous reviewer for comments on an earlier version of this paper.

about how these components interact and how, in spite of this complex interaction, the concept of a modular grammar can be preserved.

For reasons of space, the data and the resulting analysis have been reduced to the facts vital for the discussion in this paper. For an extended version, see Bögel (forthcoming).

## 2 Architectural assumptions

It is very important to understand that this paper is not written from the perspective of a grammar where all modules “are present in parallel” (Asudeh and Toivonen 2010). Instead, the underlying assumption is that any act of language necessarily takes a certain direction, i.e., from meaning to form or vice versa. The analysis proposed here is conducted from the perspective of language production (human or machine) and thus from a specific *direction*: From c-structure to p-structure, i.e., from the viewpoint of generation/production. The ‘order’ of the modules follows findings on the arrangement of grammatical modules from the field of psycholinguistics.<sup>1</sup> As this notion of a *pipeline* architecture seems to contrast with the notion of the *parallel* architecture, the position taken in this paper is explained in the following.

In the approach presented here, the notion of “parallel modules” refers to the level of *competence* in that each module and the constraints and principles related with it are stored in parallel. Furthermore, the connections between these modules, the correspondence functions, are stored along as well. Together, they form the individual’s (static) knowledge of a language, its grammar. Note, however, that projection functions always imply a direction; they are part of the interface between two modules A and B and can be understood as the output, e.g., module A feeds into module B (or as the reference to a constraint stored in the static knowledge of module B). While the *knowledge* about these modules is present in parallel, the projections between them have to be triggered by a specific process of module A before they can project their information to module B. Projection functions are thus necessarily directional, an assumption also supported by the fact that they are not reversible and are thus difficult to grasp in the context of ‘parallelity’.

As stated above, any act of language is directional. This means that as soon as a specific linguistic phenomenon is analysed, the analysis necessarily takes a direction — this is especially true if several modules are involved. Such an analysis does *not* fall under the notion of *performance* as defined by Chomsky (although it can be interfaced with it): The analysis presented in this paper is not concerned with individual speaker circumstances or memory constraints, but with the abstract rules

---

<sup>1</sup>Note that researchers in the field of psycholinguistics largely agree on the order of the modules; the discussions are mainly concerned with if and how the modules overlap (see Levelt (1999) for an overview and important references). Furthermore, speech production and perception are not viewed as identical processes, i.e., speech perception is *not* the reverse of speech production. This assumption is, in fact, nicely accounted for by the non-bidirectionality of correspondence functions in LFG and stresses the need for a directional analysis.

and constraints stored in each language module that are necessary for the analysis or production of a specific phenomena. The application *tests* the constraints stored in each module and the processes at the interface on an abstract level and from a specific direction. If the application fails, then the *competence* level, i.e., the constraints and rules stored in each module, needs correction.

It is, of course, possible to avoid such a pipeline view and analyse a linguistic phenomenon in a continued parallel view of the modules. However, it is unclear how such an analysis contributes to the overall goal of linguistic analysis which is to add knowledge to the level of competence on the basis of a concrete application in a specific direction, and to test the resulting competence level by further applications. The 'static' (parallel) analysis of a phenomenon, especially of one involving separate modules, does not contribute to this process as it will necessarily fail in a concrete case of linguistic application (machine *and* human language processing).

### 3 Degema – some background

Degema is a Delta-Edoid language, spoken in the Rivers State region of Southern Nigeria. So far, no standard version of Degema has emerged, but there are two dialects: Usokun and Atala, spoken by an estimated 11.000 speakers each. The main focus of this paper lies on the Usokun dialect. From the phonological viewpoint, Degema has several interesting aspects related to the topic of this paper.

#### 1. Syllable structure:

Degema has four basic syllable types: V, VC, CV and CVC. Consonant clusters are the result of deleting an intervening vowel, thus CCV is derived from CVCV. In connected speech, consonants will resyllabify to a following word starting with a vowel. Thus, the prevalent syllable structure is CV or CVC (cf. Kari 2004, 378ff.).

#### 2. Vowel harmony:

The ten Degema vowels can be divided into two sets: One set with an advanced tongue root (phonetic feature +ATR) and one set with a retracted tongue root (-ATR) (Fulop et al. 1998). Vowels in simple words are exclusively drawn from one of the sets. Vowel harmony also spreads to clitics, possessives, some object pronouns and the negative adverb (Kari 2007).

#### 3. Lexical tone:

Degema is a tone language, which can distinguish between segmentally and categorically identical lexical entries via lexical tone and thus requires an elaborate analysis of all linguistic aspects of a lexical entry. Degema has two basic tones, H(igh) (x̂), L(ow) (x̃, or unmarked) and a downstepped high tone (indicated by ↓x), which applies if a High tone directly follows another High tone. Tones are marked on the syllables of a string. As reported by Kari (2004), there is no evidence of contour tones, i.e., two tones do not share one

syllabic host. Sequences of tones occur only in a handful of words. These consist of identical vowels and have a high - downstepped high tone pattern.<sup>2</sup>

## 4 The factative clitic

Degema has several clitics; of these, the factative clitic is of special interest, as it can appear as an enclitic in some contexts, but as an endoclititic in other contexts. The factative clitic attaches to verbs and object pronouns. The factative aspect

*is used to denote a fact, which may be a dynamic situation that has already been completed or a state that once existed or still exists at the present time.* (cf. Rose 2014)<sup>3</sup>

The factative clitic has an underspecified vowel, which copies the features of the vowel in the host's last syllable as part of a process of vowel harmony, followed by an *n*: *Vn*. The clitic's realisation depends on

- a. the phonological environment (consonant vs. vowel) of the host's last segment and (if present) the following word's first segment.
- b. its medial vs. final position in an intonational phrase (IntP).

In the following sections, the different realisations of the clitic will be demonstrated with a number of (non-IPA) examples from Kari (2004).<sup>4</sup>

### 4.1 Enclisis in the intonational phrase (IntP) medial position

In the medial IntP position, the underspecified vowel is never realised. The realisation of the *n* is determined by its phonological environment (FE  $\hat{=}$  factative).

#### 1. Phonological environment: **Vowel–Factative–Consonant (xxV=n Cxx)**

- (1) Breno o=síré t́á=**n** mú éki  
 Breno 3Sg=run go=FE to market  
 'Breno ran to the market.' (Kari 2004, 114)

#### 2. Phonological environment: **Vowel–Factative–Vowel (xxV=n Vxx)**

- (2) Ení ðól-ám ójzí yọ i=đíyómósé=**n** ávom ðáaw  
 we hold-GER thief DEF 3Sg=sweeten=FE inside their  
 'It pleased them that we caught the thief.' (Kari 2004, 50)

<sup>2</sup>There are no long vowels in Degema (the word for 'yes' being an exception).

<sup>3</sup>Kari prefers the notion *factative*, because it "marks past in dynamic verbs but past/non-past in stative verbs. Given this situation, one can really not describe factative as perfective, since in stative verbs factative could have a non-past or timeless meaning/interpretation." (Rose 2014, fn 4).

<sup>4</sup>Note also that while the occurrence of the factative clitic is quite common, it never occurs in (sub)clauses with a negative meaning. This aspect and its implementation is 'ignored' as it involves the semantic component as well – an analysis would thus go far beyond the scope of this paper.

3. Phonological environment: **Consonant–Factative–Vowel (xxC=n Vxx)**

- (3) Uḅuwan i=kél=**n**                      úsóm yọ  
salt      3Sg=be more than=FE soup DEF  
‘Salt is more than the soup.’                      (Kari 2004, 153)

4. Phonological environment: **Consonant–Factative–Consonant (xxC ∅ Cxx)**

- (4) È=yáw      mú      ínwíny      útany  
3Pl=take.FE from body tree  
‘They got it from a tree.’                      (Kari 2004, 200)

In this specific environment, the factative enclitic is not realised which is consistent with the general rejection of complex consonant clusters mentioned in Section 3. Thus, the generalisation can be established that the factative clitic is realised as *n* in medial IntP position, except when preceded and followed by a consonant.

## 4.2 En(do)clisis at the IntP final position

In the case of the factative clitic *Vn* following a vowel at the final IntP position, the underspecified vowel copies the features of the host’s last vowel.

1. Phonological environment: **Vowel–Factative (xxV=Vn)**

- (5) O=síré=**↓en**  
3Sg=run=FE  
‘(S)he ran.’                      (Kari 2004, 72)

Note that there is no simple vowel lengthening, as Degema does not feature long vowels and does not allow for contour tones (see Section 3). Instead, the presence of the factative vowel is indicated via a downstepped High tone.

2. Phonological environment: **Consonant–Factative (xxVVC) → endoclisis**

- (6) O=ḅó<sup>↓</sup>ol  
3Sg=hold.FE  
‘(S)he held (a cloth).’                      (Kari 2004, 72)

As in (5), the underspecified vowel copies all the features of the host’s last vowel; it is then moved into the last syllable via *metathesis* (by switching position with the last consonant of the host) (Kari 2002). However, in order to avoid illegal consonant clustering, the *n* is deleted (which is consistent with the deletion in the context *xxC ∅ Cxx*, see (4)). The result is a changed tone pattern H<sup>↓</sup>H as found in (5), which contrasts with the factative clitic in IntP-medial position as shown in the parallel construction in (7).

- (7) O=ɸól=n ítónw  
 3Sg=hold=FE cloth  
 ‘(S)he held a cloth.’ (Kari 2004, 72)

A possible alternative explanation would be that the factative clitic in IntP-final position causes the vowel of its host to be lengthened and to adapt two different tones. However, this is not conform with the general phonological findings of Degema, which do not allow for vowel lengthening or contour tones.

The following figure shows a Praat (Boersma and Weenink 2013) analysis of a sentence containing endo- and enclisis.

- (8) [ómóβítám jɔ ɔ=sɔ=↓ɔ=l ] [ɔ=kírí béné=n úkpá]  
 girl DEF 3Sg=jump.FE 3Sg=also play=FE dance  
 ‘The girl jumped and danced.’ (Kari, p.c.)

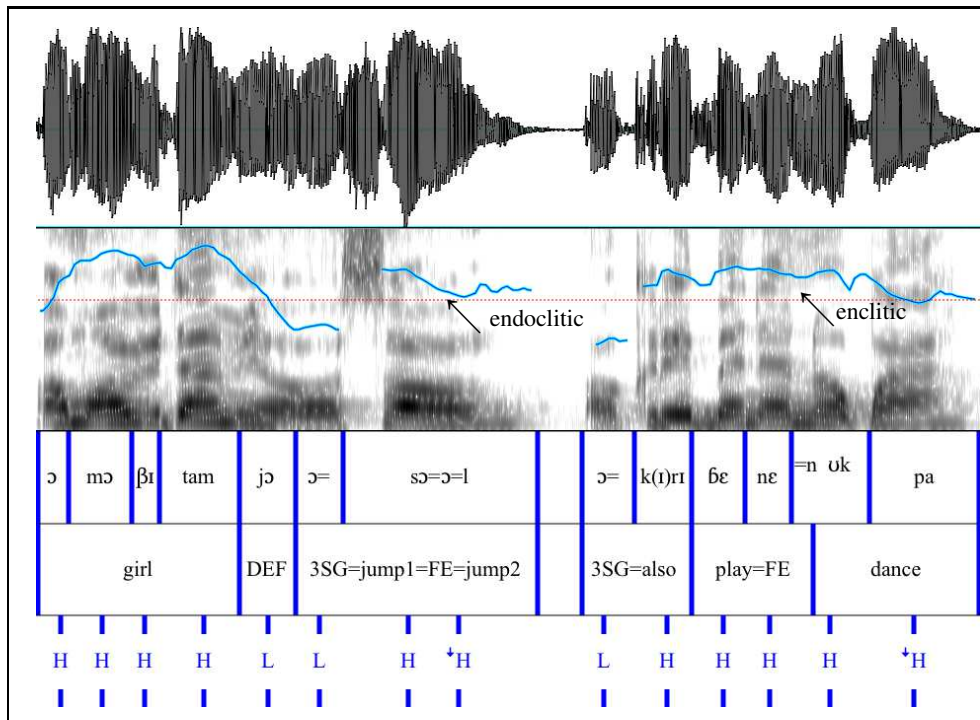


Figure 1: Speech signal for (8): ‘The girl jumped and danced.’

Figure 1 consists of two IntP. The first IntP is terminated by the factative clitic. Since the host ends in a consonant, the clitic moves into the host and the *n* is deleted. The clitic’s vowel is marked by a downstepped High tone, as indicated by the first arrow in the speech signal, where the pitch clearly moves from a high pitch mark to a lower one.

The second part of the sentence contains a factative clitic in a medial IntP position. Here, the vowel is not realised and no downstep occurs; the factative clitic is represented by an *n*, as indicated by the second arrow in the speech signal.

### 4.3 Prosodic phrase or syntactic clause

As mentioned before, the sentence in Figure 1 can be prosodically divided into two intonational phrases (IntP). This division is isomorphic to the syntactic analysis of the sentence in that each IntP matches a larger syntactic phrase. And indeed, as the following example shows, the realisation of the phrase-final en(do)clitic is not only restricted to the sentence-final IntP position, but also occurs at the final position of a subordinate clause and/or a matrix clause ((9)) forming an IntP.

- (9) [E=kótú mé=**en**] dọ mí=meme  
3Pl=call me=FE but 1Sg.NEG=answer  
'They called me but I didn't answer.' (Kari 2004, 139)

These examples raise another question: If the intonational phrase boundary and the syntactic clause boundary are isomorphic then which module (syntactic or prosodic) is responsible for triggering the factative endoclititic? Evidence for the intonational phrase boundary as a trigger for the factative endoclititic comes from the definite marker *yọ* which usually appears directly behind the noun. However, this sequence can be interrupted by a subordinate clause. In this case, the marker can appear behind the factative clitic ((10), round brackets indicate prosodic phrasing and square brackets indicate syntactic phrasing).

- (10) Mì=món owéy ([nú (baw) e=kótú=**n**] yọ)  
1Sg=see=FE person that they 3Pl=call=FE DEF  
'I saw the person who they called.' (Kari 2004, 53)

In the above example, the definite marker *yọ* is phrased into the same intonational phrase as the clitic. Syntactically, however, the definite marker is external, as it belongs to the 'person' of the sentence. If the form of the clitic was determined syntactically, then the endoclititic should be triggered in (10). Instead, the clitic is realized as if in phrase medial position which follows from the fact that the clitic indeed is not final if considering not the syntactic, but the prosodic phrasing. Following its paradigm for the medial IntP position between a vowel and a consonant, the factative is thus realised as the enclitic *n* in (10).

## 5 Issues caused by endocclisis

The concept of endocclisis poses a challenge to linguistic theory. As it involves several modules of language it is the optimal candidate for interface discussions. In the following sections, two of the issues are discussed in detail.

### 5.1 Lexical Integrity

One of the seemingly difficult problems is the violation of the concept of Lexical Integrity, as stated in Bresnan (2001, 92):



(11) **Lexical Integrity:**

Morphologically complete words are leaves of the c-structure tree and each leaf corresponds to one and only one c-structure node.

Clitics are independent syntactic items which occupy a separate terminal node (Butt and King 2005) and should not be able to disrupt a morphologically complete word. Endoclitics move into the word (stem), which is why they have been named a challenge to the principle of Lexical Integrity, especially to the Strong Lexicalist Hypothesis which states that rules of phrase formation cannot access any part of a given word.<sup>5</sup>

However, one should carefully distinguish between the different components involved in the generation of endoclitics. The principle of Lexical Integrity is concerned with the interaction between syntax and morphology and the separation between words and phrases. The Degema clitics, on the other hand, also involve postlexical phonology in p-structure and it is the discussion of this component and the resulting architectural decisions that will allow for the realization that lexical integrity and endoclitisis are not exclusionary concepts, but, on the contrary, are complementary from a specific architectural viewpoint.

## 5.2 Modularity

The previous section briefly addressed the fact that the analysis of endoclitics requires a theory of interfaces between different modules of grammar. In the case of Degema endoclitics, the analysis requires an interaction of several modules/components of grammar: c-structure, p-structure and postlexical phonology, and the lexicon. The involvement of several modules requires careful definition of boundaries and responsibilities, because it is commonly assumed that one module cannot refer to or process vocabulary of another module in the grammar.

*From a formal point of view, the hypothesis of [...] Modularity claims that the informational architecture of the mind strictly segregates phonological, syntactic, and conceptual representations from each other. Each lives in its own module; there can be no 'mixed' representations that are partly phonological and partly syntactic, or partly syntactic and partly semantic. Rather, all coordination among these representations is encoded in correspondence rules.* (Jackendoff 1997, 83)

This means that, e.g., postlexical phonological rules cannot refer to syntactic boundaries, because they do not understand about syntactic boundaries. To enable communication between modules, the output of one module has to be 'translated' into the vocabulary of the receiving module. Only then can information be processed.

While Jackendoff's approach is slightly different from the approach presented here, the above quote fits nicely into the architectural ideas and concepts of LFG.

---

<sup>5</sup>See Asudeh et al. (2013) for a thorough discussion of the concept of Lexical Integrity in LFG.

Correspondence rules project (and ‘translate’) information between modules, thus preserving each module’s inherent processes and vocabulary, while at the same time allowing for a correspondence between the different modules by means of projections. The challenge here lies in the question as to how the information on syntactic structure should be ‘translated’ into prosodic information that can be referred to by the postlexical phonological rules, especially under the assumption that prosodic and syntactic constituent boundaries can be mismatched (see, e.g., Nespor and Vogel (1986)). However, before a solution can be discussed, some basic assumptions about the string in LFG have to be established.

## 6 Assumptions about “the string”

Discussions about modularity, lexical integrity and endoclysis tend to involve a discussion about the nature of the string as well. Dalrymple and Mycock (2011) proposed a distinction between p(honological)- and s(yntactic)-string as two sides of the string. They place this at the heart of the prosody-syntax interface, with the lexicon as a reference to match the two sides of the string. While a set of (lexical) phonological rules apply between p-forms and p-string, (postlexical) phonological rules, whose domains are the higher prosodic units, are assumed to apply between p-string and p-structure. Prosodically triggered phenomena like Degema endoclysis would thus be part of the relation between p-string and p-structure while the parallelity of p-string and s-string would be preserved.<sup>6</sup>

The following sections pursue the question if the linear order of the string that we perceive in listening is equal to the linear order of the string that we analyse syntactically; or if we put it into the terms coined by Dalrymple and Mycock (2011): Is the p(honological)-string parallel to the s(yntactic)-string? Or, as a final perspective: Are the postlexical phonological rules which operate on the postlexical string a translator between s- and p-string or are they an external interpreter needed for the recognition of lexical elements (thus leaving the parallelity of s- and p-string intact)? Each of the options for the postlexical phonological rules component has consequences for the nature of the string interface. The following sections give an overview by discussing the two possible architectural assumptions from the viewpoint of perception/parsing via the following example, where the factative clitic is attached phrase-finally to an object pronoun (*báw*).

- (12) ɔ= nú bá<sup>↓</sup>aw  
 3Sg= hit them.FE  
 ‘S/he hit them.’ (Kari 2002, 45)

The conclusion that the syntactic and the phonological string are non-parallel implies specific assumptions about the grammar architecture whose concrete realisation will be shown in Section 7.

<sup>6</sup>However, as Dalrymple and Mycock (2011) do not focus on this particular area, an exact elaboration of these processes and the resulting issues within their framework is part of future work.

## 6.1 Postlexical phonological rules in a translation related function



Figure 2: Partial syntactic tree and architecture for s-string  $\neq$  p-string.

In Figure 2, the s-string is not parallel to the p-string; the clitic occupies its own terminal node, following its host in an *enclitic* position. The postlexical phonological rules component, which is ultimately responsible for the positioning of the clitic within the host, is situated between the s- and the p-string and has a translation related function in that it decomposes the incoming speech signal before the lexical analysis according to the language's inherent postlexical phonological rules. As a result, p- and s-string are parallel most of the time, but might differ in the case of phonological intervention (endoclysis, second position clitics). From this perspective, the principle of lexical integrity is not violated, as the clitic occupies its own terminal node. Furthermore, an analysis in general is much easier, as the clitic is not locked into another syntactic (and possibly completely unrelated) element (e.g., the pronoun). However, there are also drawbacks to this view, e.g., if the 'original' position of the clitic is unknown, as it is the case with some second position clitics: If they do not have a corresponding full form, it is difficult to motivate their 'movement' through the application of postlexical phonological rules.

## 6.2 Postlexical phonological rules in a interpreting function

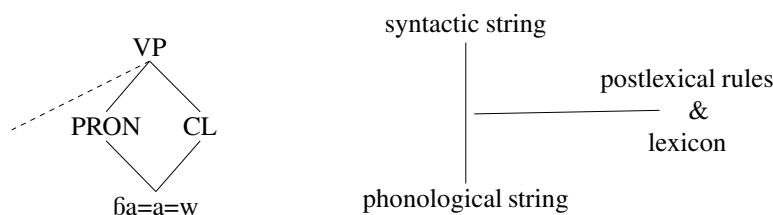


Figure 3: Partial syntactic tree and architecture for s-string = p-string.

Under this view, the parallelity between p- and s-string is preserved. The postlexical component is an external resource which is used together with the lexicon to align the speech signal with the syntax. And although the clitic is located (locked) within another syntactic element, an analysis could be possible with the application of inside-out functional uncertainty (e.g. Nordlinger 1998).

However, there are two serious issues with this approach: First, the concept of two terminal nodes sharing one lexical item is a violation of the principle of Lexical Integrity.<sup>7</sup> Second, there is also a problem from the perspective of modularity. The analysis under one terminal node can only be explained from the perspective of *perception/parsing*, because only then is the information on prosodic phrasing available before lexical and syntactic analysis; i.e., the trigger for endocclisis is available before the clitic and its host are analysed by syntax. However, it is difficult to explain how the endoclititic came to be within the pronoun if the opposite direction (*production/generation*) is considered. In this case, the trigger for the process of endocclisis, the intonational phrase boundary, is ‘not available’ as of yet, but will only be available at a ‘later’ stage.

Both of these issues are unresolvable under the current assumptions. It can thus be concluded that the postlexical phonological rules have a translation related function and that while the s- and p-string are isomorphic most of the time, under certain circumstances, they do not have to be. The following section formalizes this approach and shows how modularity and lexical integrity can be maintained through the proposed analysis of endocclisis on the basis of two assumptions:

1. The linear order of the “speech signal” (the p-string) is not necessarily equal to the linear order of the s(yntactic)-string
2. Assuming modularity, the postlexical phonological rule responsible for endocclisis refers to prosodic (and not syntactic) boundaries.

## 7 Endocclisis and the syntax-prosody interface

Figure 4 gives an overall overview of the assumed architecture applied to example (12).<sup>8</sup>

---

<sup>7</sup>At this point it is worth discussing why the approach presented in this paper is to be preferred to the Lexical Sharing approach presented by Wescoat, specifically relating to his analysis of Udi endocclisis (see Wescoat (2009) and references therein). First of all, the clitic types discussed in the two papers are of a different nature. In contrast to the Degema factative clitic, Udi (endo)clitics are not triggered by prosody. Second, Lexical Sharing cannot explain per se the occurrence of endocclisis. The approach has to rely on additional formal power (in the form of Optimality Theory) to account for these phenomena. Third, it has been shown in Bögel (2010) that the treatment of clitics within the lexicon, as proposed by Lexical Sharing, results in a listing of possible combinations, which, in the case of promiscuous clitics, can be quite extensive. This potentially infinite enlargement of the lexicon is replaced by a (finite) set of abstract rules accounting for endocclisis in p-structure. Fourth, the approach presented here keeps the principle of Lexical Integrity intact. In contrast, Lexical Sharing relies on a modified version thereof, the *Homomorphic* Lexical Integrity which states that the precedence relation between two terminal nodes must be identical to the precedence relation of the corresponding lexical exponents, thus allowing for a one-to-many relation between the lexical exponent and the corresponding terminal node(s) (strictly speaking, the precedence relation is no longer given with endocclisis). It can thus be concluded that, in the case of prosodic clitics the approach presented here should be preferred, because it respects the concept of Lexical Integrity and the concept of modularity while at the same time explaining a complex phenomena without the application of additional formal power outside of the grammatical modules assumed in linguistic analysis.

<sup>8</sup>For space reasons, Figure 4 is reduced to the information needed for the analysis of en(do)clisis.

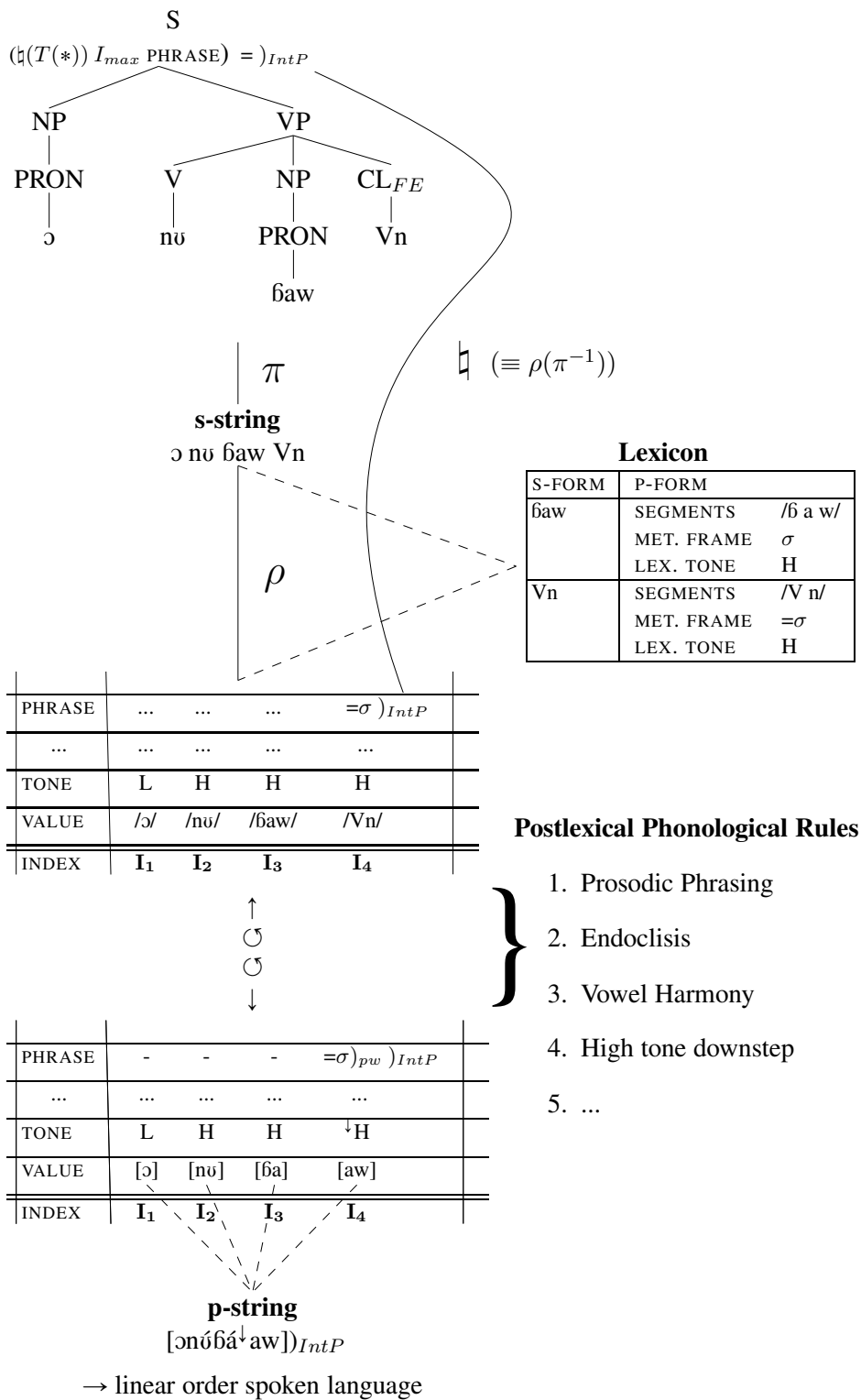


Figure 4: The integration of prosody and postlexical phonology into LFG.

The architecture proposed in Figure 4 shows how postlexical phonology and prosody can be integrated into LFG. The upper part represents c-structure. The lower part, consisting of the two p-diagrams and the set of postlexical phonological rules, represents p-structure. In accordance with the assumptions made in the introduction about the directionality of each act of language, the ‘upper’ (henceforth ‘preliminary’) p-diagram represents the *input* to the p-structure module, consisting of information from two other grammar components: syntax and lexicon.<sup>9</sup>

The *postlexical phonological rules* are the constraints and rules stored in the p-structure module. They apply to the input (the information stored in the preliminary p-diagram) in a cyclic manner (indicated by  $\cup$ ), as the output of one rule often creates the context that triggers another rule. In Degema, for example, this means that the creation of a prosodic word domain has to be completed before the process of Vowel Harmony can apply, as it is restricted to this particular domain. (Postlexical) phonological rules are thus members of an ordered cascade, which can include sets of rules applying in parallel, but which allows for the application of rules to contexts that have been altered by previous rules. The ‘lower’ (henceforth *final*) version of the p-diagram is the output of p-structure after the phonological rules have applied.

All three parts belong to the module of p-structure, although their description can change. In the case of perception/parsing, the above process would be reversed: The formerly *final* p-diagram is the input, the cascade of rules applies in the opposite direction, and the formerly *preliminary* p-diagram would be the output of the p-structure module. This is of course a simplified description, as not all processes of production are reversible to become processes of perception, but there are two major insights. First, the underlying concept stays the same: There is an input to which rules and constraints are applied to form an output, and second: It is vital for any linguistic analysis to state the perspective from which an analysis is conducted.

The following sections will explain the architectural assumptions behind the syntax→prosody interface and the processes in p-structure step by step: First by distinguishing between two major interface interactions: The *transfer of vocabulary* on the one hand and the *transfer of structure* on the other.<sup>10</sup> Following this will be a section of the postlexical phonological rules that are inherent to p-structure.

## 7.1 Transfer of vocabulary

*Transfer of vocabulary* refers to the transformation of syntactic items (s-forms) to phonological representations. At the heart of this process lies the multi-layered lexicon which contains the respective information and serves as a medium for the lexical look-up (Bögel 2013). The multi-layered lexicon distinguishes between three dimensions for each lexical entry (cf. Levelt et al. 1999): a (semantic) concept, a s(yntactic)-form and a p(honological)-form. The lexicon’s output is assumed to consist of (morphologically) complete words as proposed by Butt and Kaplan

---

<sup>9</sup>The integration of information from other modules is left for further research.

<sup>10</sup>These terms were coined by Scheer (2011, 558), but have a slightly different meaning here.

(2002). The lexicon as it is represented here can thus be viewed as a *lexical surface representation* (Mohan 1982). It is accessed between c- and p-structure,<sup>11</sup> aligning possible p-forms with corresponding s-forms and vice versa. The dashed lines emphasize its role as a look-up instrument. Table 1 shows the lexical entries for the factative clitic and the pronoun *ɓaw* ‘they’ (example (12)).

CONCEPT	S-FORM	P-FORM
‘they’	ɓaw PRON (↑ PRED) = ‘they’	P-FORM [ɓáw]
	...	SEGMENTS /b a w/
	...	MET. FRAME $\sigma$
	...	LEX. TONE H
<i>factative</i>	Vn CL (↑ EVENT-TYPE) = factative	P-FORM [V̂n]
	...	SEGMENTS /V n/
	...	MET. FRAME $=\sigma$
	...	LEX. TONE H

Table 1: Lexical entries for the factative clitic and the pronoun *ɓaw* ‘they’.

Table 1 focuses on the p-form<sup>12</sup> which has (in this case) four attributes: SEGMENTS represents the single segments of a p-form; METRICAL FRAME shows the number of syllables and, in the case of the clitic’s entry, describes the prosodic status:  $=\sigma$  indicates that the corresponding p-form is prosodically deficient and needs to lean on a host to its left. LEXICAL TONE indicates the tone attached to each syllable. Finally, P-FORM represents the element as it would be expressed in *isolation*.

In the case of a vocabulary transfer, the syntactically tokenized s-string is divided into s-forms which are matched with the related p-form in the lexical entry. The information gathered in the p-form is then projected syllable by syllable via the relation  $\rho$  to the (preliminary) p-diagram, namely the value of the respective syllable and lexical tone.<sup>13</sup> Furthermore, *structure* up to the word level is transferred to the PHRASE section in the form of the lexical metrical frame.<sup>14</sup>

	↑	↑		↑
PHRASE	...	...	...	$=\sigma$
LEX. TONE	L	H	H	H
VALUE	/ɔ/	/nɔ/	/ɓaw/	/Vn/
INDEX	I <sub>1</sub>	I <sub>2</sub>	I <sub>3</sub>	I <sub>4</sub>

Table 2: *Transfer of vocabulary* from lexicon to p-diagram, encoding example (12).

<sup>11</sup>Note that this paper does not make a point on the exact position of the lexicon. It is assumed that the position shown in Figure 4 is one of many possible access positions to the lexicon.

<sup>12</sup>The semantic CONCEPT is only superficially represented, as it is of no interest here. The S-FORM contains all information usually connected with lexical entries in LFG (cf. Butt and Kaplan 2002).

<sup>13</sup>In contrast to the original p-diagram introduced by Bögel (2012), the index is now indicated by the capital letter I (formerly S) to avoid confusion with S=syllable and S=sentence.

<sup>14</sup>The transfer of structure below the word-level as it is shown in Table 2 is reduced to the facts relevant for the discussion. See Bögel (forthcoming) for an extended version.

## 7.2 Transfer of structure

Under the assumption that syntax influences prosody at least on the higher levels, this paper follows the proposal by Selkirk (2011) in that every syntactic clause matches an intonational phrase (IntP) and every syntactic phrase matches a phonological phrase (PhP). This means that if there is a CP/S in c-structure, then there will be an (IntP) border indication in the p-diagram and so on.<sup>15</sup> This information is projected from the respective syntactic nodes via the following annotation pattern:

$$\begin{array}{c} \text{S} \\ (\natural(T(*)) I_{max} \text{ PHRASE}) = )_{IntP} \end{array}$$

The relation  $\natural$  is defined as  $\rho(\pi^{-1})$  (cf. Bögel 2013) and directly projects information from c-structure to the p-diagram. This annotation can thus be read in the following way: Consider all the terminal nodes of the current node (T(\*)) (i.e., S). From these nodes take the syllable with the maximum I index ( $I_{max}=I_4$  in the p-diagram). Add the value  $)_{IntP}$  (a right IntP boundary) to the attribute PHRASE at this specific index position. The result is shown in Table 3, where the information about the higher prosodic units is added to the already present prosodic units below the word-level.

PHRASE	...	...	...	= $\sigma$ ) <sub>IntP</sub>
LEX. TONE	L	H	H	H
VALUE	/ɔ/	/nʊ/	/βaw/	/Vn/
INDEX	I <sub>1</sub>	I <sub>2</sub>	I <sub>3</sub>	I <sub>4</sub>

Table 3: Preliminary p-diagram, encoding *structure* and *vocabulary* of ex. (12).

This very first version of the p-diagram preserves the syntactic linear order – it is mainly a ‘translation’ of syntactic terms into phonological terms. The preliminary version of the p-diagram in Table 3 allows for a reference to segments and their respective features as well as to an interim version of prosodic phrasing. However, missing is an explanation of endoclysis and other postlexical processes. These final adjustments are accomplished p-structure internally via a set of postlexical phonological rules operating on the preliminary version of the p-diagram.

## 7.3 The postlexical phonological rules component

As the name already suggests, postlexical phonological rules operate on a postlexical level, thus following the assumption that there are two levels where phonological rules apply: Lexical and postlexical (Kiparsky 1982, Lahiri 2000, a.o.) where

<sup>15</sup>This is in principle only a translation of inherently syntactic terms which has been criticised by, e.g., Scheer (2011). However, there is no denying syntactic influence on prosodic formation and it will soon become clear how these ‘mere translations’ are transformed into ‘real’ prosodic units.



the latter is taken to be (partly) responsible for a mismatch between the p-string and the s-string. In Figure 4, the postlexical phonological rule component consists of a set of ordered rules operating on the preliminary p-diagram derived by the transfer of vocabulary and structure. As a first step, prosodic phrasing is adjusted, for example, it can be assumed that the clitic forms a prosodic word together with its host. Adjustments can also be carried out at higher prosodic levels, thus accounting for non-isomorphism between syntactic and prosodic boundaries (see (13) below). Other postlexical processes operate on the segmental level, e.g., the realisation of the factative in different contexts. On the basis of the Degema data presented in Section 4, the following paradigm can be derived:

phrase position	phonological environment
medial	$xxV=n Cxxx$
medial	$xxV=n Vxxx$
medial	$xxC=n Vxx$
medial	$xxC \emptyset Cxxx$
final	$xV\downarrow VC$
final	$xV=\downarrow Vn$

Table 4: Paradigm of the factative clitic in Degema.

This paradigm can be realised by the following constraints and processes in the order presented here (where  $)_{pw}$  stands for a right prosodic word boundary,  $V$  stands for Vowel,  $C$  for Consonant and  $v$  for an underspecified vowel).

1. if clitic present then incorporate into prosodic word domain of host:  
 $\dots)_{pw} =\sigma \longrightarrow \dots =\sigma)_{pw}$  (can be applied repeatedly)
2. if factative in IntP medial position, then realise as  $n$ ; delete in context  $C \_ C$ :  
 $=vn \longrightarrow n / C \_ V, V \_ C, V \_ V$  and  $=vn \longrightarrow \emptyset / C \_ C$
3. if factative in IntP final position, then realise as  $vn$ :  
 $=vn \longrightarrow =vn / [\dots] \_ )_{IntP}$ 
  - if host ends in  $C$ , then swap position with  $C$  and delete  $n$ :  
 $Cvn \longrightarrow vC / \_ )_{IntP}$
4. apply vowel harmony:  $v \longrightarrow V_i / V_i \_ C)_{IntP}$
5. apply tone downstepping:  $[+H][+H] \longrightarrow [+H][\downarrow H]$

Applied in this order to the preliminary p-diagram in Table 3, the application of these postlexical processes presented here derives the final p-diagram for example (12), which can be viewed as the combined information provided by syntax, lexicon and postlexical phonological rules to the speech-signal-in-production.

PHRASE	...	...	...	$\sigma)_{pw})_{IntP}$
...	...	...	...	...
LEX. TONE	L	H	H	↓H
VALUE	[ɔ]	[nʊ]	[ʁa]	[aw]
INDEX	I <sub>1</sub>	I <sub>2</sub>	I <sub>3</sub>	I <sub>4</sub>

Table 5: The final p-diagram of of example (12).

#### 7.4 A note on Perception/Parsing and Production/Generation

Architectural assumptions should always consider both: parsing/perception and generation/production. The above analysis described the interface from the generation/production side. This section will quickly show some of the differences that arise if both directions, parsing and generation, are considered.

The first difference affects the representation of the p-diagram: Depending on generation or parsing, the representation changes slightly: In parsing, very concrete facts from the speech signal itself can be encoded, e.g. the Hertz values of the pitch, but not lexical information (e.g., lexical tone). From the perspective of generation, this data from the speech signal is not available. Thus, the final p-diagram from the viewpoint of generation is the combined information provided by syntax, the lexicon and postlexical phonological rules while the p-diagram from the viewpoint of parsing is rather a representation and interpretation of the speech signal itself.

Another difference is concerned with the reversibility of postlexical phonological processes. In generation, the *transfer of vocabulary* is similar to the transfer in parsing: The interaction is mediated via the multi-layered lexicon in both directions; postlexical phonological processes like the formation of an endoclititic are reversible because the corresponding rule can be applied backwards, separating the two lexical items so they can be matched against the lexicon.

In contrast, the *transfer of structure* differs between generation and parsing. From the perspective of generation, the transfer of structure gives a first indication of prosodic phrasing, which is retained, if syntax and prosody happen to be isomorphic. However, in cases like (10), repeated in (13), where the two components are non-isomorphic, prosodic phrasing is adjusted as part of the postlexical phonological processes.

- (13) Owéy ([nú ábo i=vúwóy] yo)  
 person that hands 3Pl=be dirty.FE DEF  
 ‘The person whose hands are dirty.’ (modified Kari 2004, 202)

The transfer of structure applied to (13) results in two nested Intonational Phrases, as the overall sentence as well as the embedded clause project an IntP to the preliminary p-diagram:  $i=vúwóy)_{IntP} yo)_{IntP}$ . However, in the corresponding speech signal,  $y_0$  is phrased together with the preceding IntP of the embedded clause.

It can thus be assumed that in cases, where the outer IntP of a nested structure contains only a small amount of material to the right of an embedded IntP, then the outer IntP boundary is deleted and the material is phrased together with the preceding (internal) IntP. The following table exemplifies this process.

Transfer of structure: Every CP/S projects an IntP	$(([\text{Owéy} ([\text{nú ábo i=vúwóy}]_{CP})_{IntP} \text{yo}]_S)_{IntP})$
Prosodic Rephrasing: Adjustment of IntP	$(([\text{Owéy} [\text{nú ábo i=vúwóy}]_{CP} \text{yo}]_S)_{IntP})$

Table 6: Structure transfer and prosodic rephrasing in cases of non-isomorphism.

The prosodic rephrasing is not ‘reversible’ if coming from the perspective of parsing. In the speech signal, there is no indication of an IntP boundary before the definite marker *yo*. The information that a CP boundary is present at this position is only available at a ‘later stage’.

As a consequence, it can be assumed that prosodic phrasing is not necessarily vital for the parsing of syntactic constituents – i.e., a missing IntP boundary after the CP in example (13) will not result in a failure of syntactic phrasing, albeit prosodic phrasing *can* reflect syntactic phrasing and aid in the disambiguation of syntactic ambiguities (see, e.g., Bögel (2013)).

## 8 Conclusion

This paper showed how the linguistic issues caused by endoclisis (namely the violation of Lexical Integrity and modularity) can be resolved by the introduction of a postlexical phonological rules component inherent to p-structure. These rules operate on a preliminary version of the p-diagram, derived via a) the *transfer of structure* (c-structure to prosodic phrasing) and b) the *transfer of vocabulary* via the multi-layered lexicon. Combined, these transfer processes form the c-structure–p-structure interface (or more commonly phrased: the syntax–prosody interface). The output, the ‘final p-diagram’, is the contribution of syntax, the lexicon and postlexical phonology to the speech signal.

It has also been shown that it is vital for the analysis to take a specific directional perspective: From meaning to form or vice versa. Furthermore, the opposite direction should be considered as well during the analysis in order to uncover wrong conclusions that are not identifiable as such from a monodirectional perspective. The resulting architectural assumptions allow for an analysis of endoclisis independently of syntax; thus, the modularity of the different components is preserved and the concept of Lexical Integrity is kept intact.

Returning to the initial question as to how the s- and p-string should be defined and how they relate to each other, it can be concluded that s-string and p-string are parallel most of the time, but they do not necessarily have to be. In cases of endoclisis (or, e.g., second position clitics), the linear order of the p-string may change. The s-string is thus defined as the *linear order of elements as they*

would be syntactically analysed while the p-string represents the linear order of the elements as they would be pronounced in the final p-diagram. With the architectural assumptions proposed in this paper, this difference in linear order can be easily explained and forms a stable construction for the analysis of other complex phenomena at the syntax-prosody interface.

## References

- Asudeh, Ash. 2006. Direct Compositionality and the Architecture of LFG. In M. Butt, M. Dalrymple and T. H. King (eds.), *Intelligent Linguistic Architectures: Variations on themes by Ronald M. Kaplan*, CSLI Publications.
- Asudeh, Ash. 2009. Adjacency and Locality: A Constraint-based Analysis of Complementizer-Adjacent Extraction. In *Proceedings of LFG09*, CSLI Publications.
- Asudeh, Ash, Dalrymple, Mary and Toivonen, Ida. 2013. Constructions with Lexical Integrity. *Journal of Language Modelling* 1(1), 1–54.
- Asudeh, Ash and Toivonen, Ida. 2010. Lexical-Functional Grammar. In B. Heine and H. Narrog (eds.), *The Oxford Handbook of Linguistic Analysis*, Chapter 17, pages 425–458, Oxford: Oxford University Press.
- Boersma, Paul and Weenink, David. 2013. Praat: doing phonetics by computer [Computer program], Version 5.3.56, retrieved 15 September 2013 from <http://www.praat.org/>.
- Bögel, Tina. 2010. Pashto (Endo)clitics in a Parallel Architecture. In *Proceedings of LFG10*, CSLI Publications.
- Bögel, Tina. 2012. The P-diagram – A Syllable-based Approach to P-structure. In *Proceedings of LFG12*, CSLI Publications.
- Bögel, Tina. 2013. A Prosodic Resolution of German Case Ambiguities. In *Proceedings of LFG13*, CSLI Publications.
- Bögel, Tina. forthcoming. *The Prosody-Syntax Interface in LFG*. Ph. D.thesis, University of Konstanz.
- Bögel, Tina, Butt, Miriam, Kaplan, Ronald M., King, Tracy Holloway and Maxwell III., John T. 2009. Prosodic Phonology in LFG: A New Proposal. In *Proceedings of LFG09*, CSLI Publications.
- Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Blackwell.
- Butt, Miriam and Kaplan, Ronald M. 2002. ‘The Morphology-Syntax Interface’, Handout presented at the Lexical Functional Grammar Conference, Athens, Greece.
- Butt, Miriam and King, Tracy Holloway. 2005. The Status of Case. In V. Dayal and A. Mahajan (eds.), *Clause Structure in South Asian Languages*, pages 153–198, Kluwer Academic Publishers.
- Dalrymple, Mary. 2001. *Lexical Functional Grammar*, volume 34 of *Syntax and Semantics*. Academic Press.

- Dalrymple, Mary and Mycock, Louise. 2011. The Prosody-Semantics Interface. In *Proceedings of LFG2011*, CSLI Publications.
- Fulop, Sean, Kari, Ethelbert E. and Ladefoged, Peter. 1998. An Acoustic Study of the Tongue Root Contrast in Degema Vowels. *Phonetica* 55, 80–98.
- Halpern, Aaron. 1995. *On the Placement and Morphology of Clitics*. CSLI Publications.
- Jackendoff, Ray. 1997. *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.
- Kaplan, Ronald M. 1987. Three Seductions of Computational Psycholinguistics. In P. Whitelock, M. McGee Wood, H. L. Somers, R. Johnson and P. Bennett (eds.), *Linguistic Theory and Computer Applications*, pages 149–181, London: Academic Press.
- Kari, Ethelbert E. 2002. On Endoclitics: Some facts from Degema. *Journal of Asian and African Studies* 63.
- Kari, Ethelbert E. 2004. *A Reference Grammar of Degema*. Köln: Köppe.
- Kari, Ethelbert E. 2007. Vowel Harmony in Degema, Nigeria. *African Study Monographs* 28(2), 87–97.
- Kiparsky, Paul. 1982. Lexical Morphology and Phonology. In *Linguistics in the Morning Calm*, pages 3–91, Seoul, Korea: Hanshin Publishing Company.
- Lahiri, Aditi. 2000. Phonology: Structure, Representation, and Process. In L. Wheeldon (ed.), *Aspects of Language Production*, Chapter 7, pages 165–226, Psychology Press.
- Levelt, Willem J.M. 1999. Models of word production. *Trends in Cognitive Sciences* 3(6), 223–232.
- Levelt, Willem J.M., Roelofs, Ardi and Meyer, Antje S. 1999. A Theory of Lexical Access in Speech Production. *Behavioral and Brain Sciences* 22, 1–75.
- Mohanan, Karuvannur Puthanveetil. 1982. *Lexical Phonology*. Ph.D.thesis, Massachusetts Institute of Technology.
- Nespor, Marina and Vogel, Irene. 1986. *Prosodic Phonology*. Dordrecht: Foris.
- Nordlinger, Rachel. 1998. *Constructive Case: Evidence from Australian Languages*. Dissertations in Linguistics, Stanford, CA: CSLI Publications.
- Rose, Sarah. 2014. Degema. In D. Nurse, J. Hewson and S. Rose (eds.), *Verbal Categories in Niger-Congo*, Unpublished. <http://www.mun.ca/linguistics/nicol/>.
- Scheer, Tobias. 2011. *A guide to morphosyntax-phonology interface theories: How extra-phonological information is treated in phonology since Trubetzkoy's Grenzsignale*. Berlin: De Gruyter Mouton.
- Selkirk, Elisabeth O. 2011. The Syntax-Phonology Interface. In J. Goldsmith, J. Riggle and A. C. L. Yu (eds.), *The Handbook of Phonological Theory*, pages 435–484, Malden, MA: Blackwell.
- Wescoat, Michael T. 2009. Udi Person Markers and Lexical Integrity. In *Proceedings of LFG09*, CSLI Publications.