PROSODIC PHONOLOGY IN LFG: A NEW PROPOSAL

Tina Bögel†, Miriam Butt†, Ronald M. Kaplan‡,
Tracy Holloway King‡, and John T. Maxwell III*

Universität Konstanz†, Microsoft Corp.‡, PARC*

Proceedings of the LFG09 Conference

Miriam Butt and Tracy Holloway King (Editors)

2009

CSLI Publications

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

Abstract

In this paper we outline a new architecture for modeling the interaction between syntax and prosody. This architecture does not make use of correspondences between separate projections, but it is still consonant with the overall framework of LFG. We propose that prosodic information is developed in a component that operates independently of the syntax, thus allowing easy description of misalignment phenomena. We also propose a simple way of making prosodic information accessible to syntax, so that it is possible to condition syntactic rules and preferences on prosodic boundaries. We place the prosodic and syntactic components of the grammar in a pipeline configuration such that the terminal string of the syntactic tree is a sequence of lexical formatives intermixed with features inserted by the prosodic component. Depending on how they are distributed with respect to syntactic groupings, those features may or may not have an impact on the syntactic analysis.

1 Introduction

An open question in theoretical linguistics is how to characterize the interactions between the syntactic and prosodic components of a grammar.[1] One approach to this question takes syntax as primary, following the tradition of proposals made by Selkirk (1981, 1984, 1986) and Nespor and Vogel (1986) and summarized by Selkirk (2001). Under this approach prosodic information is mapped directly from syntax, and prosodic units are therefore naturally aligned with syntactic constituents. It is expected that deviations from straightforward alignment will be quite unusual, and it can become very complex and unintuitive to describe exceptions when they do appear (e.g., Cinque, 1993). The co-description architecture proposed by Butt and King (1998) and Bögel et al. (2008) uses the formal mechanisms of LFG in a concrete instantiation of this approach.[2]

In contrast, a second school of thought assumes that syntax and prosody are typically misaligned. This idea was put forward early on by scholars like Henry Sweet, Eduard Sievers, Franz Saran and Hermann Paul (Plank, 2005, see references therein). This version of the interaction has generally had little appeal to prosodic phonologists, but recent work is undertaking a reconsideration. O'Connor (2005a) and Lahiri and Plank (2009), for example, argue that a simple correspondence between prosody and syntax is more the exception than the rule. Similarly, although Mycock (2006) works within the co-description architecture, she also assumes that there is no simple correspondence between phonology and syntax.

[1]We would like to thank Ash Asudeh, Mary Dalrymple, Aditi Lahiri, Frans Plank, and Janet Grijzenhout for detailed discussions of the issues in this paper. We would also like to gratefully aknowledge funding from the DFG (Deutsche Forschungsgemeinschaft) for the Urdu *ParGram* project.

[2]The constraint-based LFG architecture is neutral between generation and recognition and so is naturally compatible with processing models of both language production and language comprehension. Traditional phonological approaches are usually biased towards the generation/production direction, describing how syntactic and semantic structures can be converted into some representation of their pronunciation. It is less obvious how the traditional approaches can be incorporated into models of comprehension.

An example of a typical misalignment is the contrast underlined in (1), discussed by Lahiri et al. (1990, 118). The substring *I talked to* groups syntactically with the prepositional phrase to its right but prosodically with the preceding noun phrase.[3]

(1) a. [[[The man] [[I] [[talked to] [in the school]]]] [is ill]]

b. ((((The man) (I talked to)) (in the school)) (is ill))

(Wheeldon and Lahiri, 1997, 357-358)

Throughout this paper, syntactic bracketing is indicated by square brackets [], while prosodic bracketing is indicated by parentheses (). Also, the bracketing shown often collapses syntactic and prosodic levels when they are unimportant for showing the grouping of constituents.

In this paper we outline a new architecture for modeling the interaction between syntax and prosody. This architecture does not make use of correspondences between separate projections but is still consonant with the overall framework of LFG. We propose that prosodic information is developed in a component that operates independently of the syntax, thus allowing easy description of misalignment phenomena. We also propose a very simple way of making prosodic information accessible to syntax, so that it is possible to condition syntactic rules and preferences on prosodic boundaries. We place the prosodic and syntactic components of the grammar in a pipeline configuration such that the terminal string of the syntactic tree (the LFG c-structure) is a sequence of lexical formatives intermixed with features inserted by the prosodic component. Depending on how they are distributed with respect to syntactic groupings, those features may or may not have an impact on the syntactic analysis.

As support for our model, we explore the prosody-syntax relationship with respect to two different types of clitic phenomena. Clitics are interesting for investigation of the prosody-syntax interface since they often reflect misalignments between prosody and syntax and therefore give us insight into what kinds of mismatches need to be accounted for (Halpern, 1995; Halpern and Zwicky, 1996).

Out of the wealth of possible clitic phenomena, we look at just two in the context of this paper. We chose these two because they have recently figured in discussions either with respect to reconsidering the prosody-syntax alignment assumption or with respect to discussions around the prosody-syntax interface in LFG. We leave aside for now a discussion of second position clitics such as those

[3]Selkirk (1995) explicitly addresses the issue of function words and their cliticization to the preceding prosodic word and takes them out of the general mapping algorithm. However, the problem is more general than just function words, as the data from Dutch below shows. Here, it is an adverb which clticizes and thus gives rise to a misalignment between prosody and syntax.

(i) Ik $(((\text{trap})_\omega$ $\text{te})_\omega)_\phi$ hard
I kick too hard
'I kick too hard.' (Wheeldon and Lahiri, 1997, 358)

found in Serbo-Croatian (e.g., O'Connor (2005a) for an analysis within LFG) and many other languages around the world.

As our first focus we examine the prosody-syntax discrepancies posed by clitics in Germanic languages. In particular, we look at Dutch pronominal clitics, which have been discussed extensively in the (mostly phonological and psycholinguistic) literature (Berendsen, 1986; Gussenhoven, 1986; Carlos Gussenhoven, 1989; Lahiri et al., 1990; Wheeldon and Lahiri, 1997, 2002). The contrast that is important for us is is illustrated in (2) (from Lahiri et al. 1990, 118).

(2) a. [[ik] [zoek [der krant]]]]
I look for her newspaper
'I look for her newspaper.' Spoken Dutch

b. ((ik zoek der) (krant))
I look for her newspaper
'I look for her newspaper.' Spoken Dutch

As can be seen, the pronominal clitic *der* is incorporated into the prosodic word to its left in (2b) rather than grouping with the syntactic constituent to its right, as in (2a), thus providing an instance of a prosody-syntax mismatch. Further discussion of this mismatch is provided in section 2.1.

Urdu *ezafe* is the second phenomenon that we examine. Bögel et al. (2008) have argued that *ezafe* is a clitic whose properties follow straighforwardly if the prosodic dimension is taken into account. They present an analysis which builds on Butt and King (1998), who implemented the interaction between Bengali clitics and prosody as analyzed by Hayes and Lahiri (1991) and Lahiri and Fitzpatrick-Cole (1999) via a *p(rosodic)-projection*. This p-projection follows the standard LFG architecture in that it is projected from the c-structure in parallel to the f-structure and thus follows the alignment assumption of Selkirk (1981, 1984, 1986), Nespor and Vogel (1986), and Truckenbrodt (1999). However, the implementation that Bögel et al. (2008) present has some difficulties which are resolved under the alternative approach presented in this paper in section 2.2.

Finally, we point to prosodically-determined resolution of syntactic ambiguities as another source of evidence for our model. Without prosodic information the string *old men and women* has two different syntactic bracketings, corresponding to two different interpretations (3):

(3) Syntactic bracketing
a. [[old men] and [women]]
b. [old [men and women]]

In our proposed pipeline architecture, the first of these would be preferred given the prosody in (4a) and the second would be preferred for the pattern in (4b):[4]

[4]We have left the bracketing of *and* somewhat underspecified in (4). It can represent a prosodic word on its own, but additional bracketings for (4) may occur when *and* is prosodically a clitic,

(4) Prosodic bracketing

a. (old men) (and women)

b. (old) (men and women)

We will use these examples to illustrate how information coming from the otherwise independent prosody component can influence the distribution of optimality-theory preferences (Frank et al., 1998) and thus affect the selection of particular syntactic analyses.

The following sections first provide more detail about the phenomena under consideration. We then introduce our LFG-oriented architecture of independent components that communicate and interact through symbols on a shared string. For the sake of concreteness, we show how the syntactic aspects of our proposal can be implemented by means of the notations and formal mechanisms that already exist in the XLE computational interpreter for LFG grammars (Crouch et al., 2009), and thus we show that this approach does not require mathematical or computational extensions of LFG syntactic theory. As a separate hypothesis, we suggest that the independent prosodic component needs no more than the mathematical and computational power of regular relations and finite-state transducers, the same devices that are already used for morphological analysis within the XLE system. Since the LFG languages are closed under pipeline composition with regular relations, and since the XLE system can perform finite-state transductions, the combination of LFG syntax with a regular prosodic component fits comfortably within the formal systems that already exist.

2 The Interaction of Prosody and Syntax

Selkirk (1986) made a particularly straightforward proposal for the interaction of prosody and syntax. She put forth the requirement that a unit of prosodic structure must have as its terminal string the stretch of the surface syntactic structure that is demarcated by the right and left ends of selected syntactic constituents. This postulates a relation of close alignment between prosodic units (inferred by their blocking or triggering of postlexical phonological and prosodic processes) and syntactic constituents (determined by traditional arguments involving substitution, coordination, extraction, and the like). The prosodic and syntactic structures are not isomorphic under this conception, because it does not require a distinct prosodic unit for every level of the syntactic hierarchy — the prosodic structure can be flatter than the syntactic. The situation where some elements of a syntactic constituent belong to one prosodic unit and other elements of that same constituent belong to another prosodic unit is then seen as being an exceptional instance of misalignment

often written *'n* in representations of colloquial English. This does not change the fact that different prosodic groupings of the prosodic words that correspond to the content words prefer different syntactic structures.

or mismatching. The items *der* and *krant* in example (2) above are thus misaligned according to Selkirk's algorithm.

Butt and King (1998) showed how the co-description architecture of LFG can be used to implement this general conception. They introduced an explicit prosodic structure (p-structure) that is projected from the c-structure by co-describing constraints in the same way that the f-structure is projected from the c-structure. Co-describing constraints can correlate selected syntactic categories to particular levels of the prosodic hierarchy, as suggested by Selkirk, and constraints that equate the prosodic units corresponding to other mothers and daughters allow for the flattening of prosodic structures, as desired. The architecture naturally permits a limited amount of "heightening" — the specification of intermediate prosodic levels that do not correspond directly to syntactic constituents — but that expressive power is not required to implement the Selkirk proposal.

Butt and King (1998) also observe, however, that the co-description architecture does not easily allow for prosodic units that are misaligned (in the sense defined above) with syntactic constituents. In this section we discuss two sets of linguistic data that suggest that misalignments are not atypical: Germanic, primarily Dutch, clitic placement and Urdu *ezafe*. Evidence of this sort is what motivates our consideration of new architectural arrangements.

2.1 Misalignment of Germanic Clitics

The Germanic languages are among those where phonological phrasing systematically diverges from syntactic phrasing (Lahiri and Plank, 2009). As a particular case in point, a series of psycholinguistic experiments has shown that the prosodic properties of Dutch clitics are misaligned with their conventional morphosyntactic properties (Lahiri et al., 1990; Wheeldon and Lahiri, 1997, 2002). The psycholinguistic reality of mismatches is demonstrated by the prosodic and syntactic phrasing of the Dutch article *de*, which, being a clitic, needs to be incorporated into another prosodic word, as in (5).

(5) Ik drink de wijn
'I drink the wine'

	ik	drink	de	wijn
	I	drink	the	wijn
Syntactic Phrasing:	[[ik]	[drink	[de	wijn]]]
Phonological Phrasing:	ik	((drink)	de)	wijn
	(Wheeldon and Lahiri, 1997, 358)			

The experiments by Wheeldon and Lahiri (1997) show that the phonological phrasing in (5) is the correct one and that the Dutch definite determiner *der/de* is indeed a clitic that forms a prosodic word with the word to its left. In one experiment, using an experimental method whereby the speaker was offered a delayed response, they sought to determine whether the number of prosodic units or the number of actual words account for the length of speaker preparation time for sentences such as the following:

(6)	a. (ik drink de) (wijn)	'I drink the wine'	(2 prosodic units, 4 words)
	b. (ik drink) (Jans) (wijn)	'I drink Jan's wine'	(3 prosodic units, 4 words)
	c. (ik drink) (wijn)	'I drink wine'	(2 prosodic units, 3 words)

By hypothesis, sentences (6a) and (6c) share the same number of prosodic units although they differ in the number of words. These should take the same speaker preparation time, while sentence (6b), which has one additional prosodic unit, should take longer for the speaker to prepare. Indeed, Lahiri and Wheeldon's results show that (6a) and (6c) do take the same amount of speaker preparation time, while (6b) takes the speaker significantly longer to prepare. This leads to the conclusion that it is the number of prosodic units and not the number of words that is relevant for speaker preparation time, and it supports the claim that the definite determiner *de* forms a prosodic word with the word to its left.

To further determine whether the definite determiner *de* in (6a) attached to the host on the left and could therefore be classified as a clitic, Lahiri and Wheeldon conducted a second experiment in which they required an immediate response from the speaker (thus allowing minimal or no time for planning). The idea behind the experiment was that it would be the *size* of the first prosodic unit that mattered for the speaker and not the number of prosodic units as in the first experiment. The result of this experiment was that sentences (6b) and (6c) took the speaker the same time to prepare, while sentence (6a) needed a significantly longer time. The conclusion drawn from this is that with the spontaneous response, the size of the first unit matters. The longer preparation time for sentence (6a) can only be explained if *de* is attached to the left and is therefore acting as a clitic.

Further proof for the Dutch definite determiner being a clitic and being integrated into the prosodic unit to the left comes from junctural rules. Consider voice assimilation in Dutch:

(7)

	expression	[gd]	[kd]	[kt]
Compound	$(\text{zak})_\omega$ $(\text{doek})_\omega$ 'handkerchief '	+	–	–
Clitic	$(\text{zoek der})_\omega$ 'seek her'	+	–	+

The Dutch compound noun in (7) shows voicing assimilation across a prosodic word boundary, with the assimilation of unvoiced *k* to voiced *g* in the context of the voiced consonant (*d*) starting the second prosodic word. In contrast, the clitic example in (7) argues for a single prosodic word because the voicing assimilation can also go the other way, from the final voiceless consonant of the host (*k*) to the initial consonant of the clitic (Gussenhoven, 1986; Lahiri et al., 1990).

This evidence shows that *de/der* are clitics in Dutch and that they incorporate into the prosodic word on their left. As our main interest is the determination of the interaction between prosody and syntax, the important observation for us is that function words such as the Dutch definite determiner syntactically group with various syntactic phrases to their right, but incorporate prosodically with various phrase types to their left. The same point is made by the English function words *a* and *of* in the natural pronunciation of the admonition in (8), discussed by Lahiri and Plank (2009).

(8)	Syntactic Phrasing	[Drink [[a pint] [of milk]] [a day]]
	Phonological Phrasing	(Drink a) (pint a) (milk a) (day)

Writing special rules to account for the variety of prosody-syntax mismatches would seriously complicate all of the prosody-syntax mapping algorithms that assume a basic prosody-syntax alignment. We therefore see the prosodic behavior of Germanic clitics as strong motivation to try for an alternative conception.

2.2 Misalignment of Urdu Ezafe

The South Asian language Urdu contains a construction traditionally called *ezafe* or *izafat*, which it borrowed via language contact from Persian. Persian *ezafe* has been analyzed in a number of papers; Samvelian (2007) provides a very thorough overview of the phenomenon and the work that has been done on it.

In Persian, the *ezafe* originated from a relative clause construction. Its interest for modern linguistics is that it does not respect the usual headedness patterns of the language and that it licenses complements from an unexpected position. Samvelian (2007), while providing an otherwise very lucid account of the properties of Persian *ezafe*, analyzes it as a morphological affix and as an instance of head-marking of grammatical relations. She does discuss the possibility of *ezafe* being a clitic but dismisses this possibility. Rather than explicitly integrating a full-blown prosodic module into her analysis, she integrates a reference to the prosodic properties of *ezafe* by including an EDGE feature by which the *ezafe* has to percolate to a clausal edge in the syntactic representation.

Urdu *ezafe* has a more restricted use than Persian *ezafe*. However, there is much of the same type of evidence as in Persian that it is a clitic and that it is prosodically incorporated into the right edge of the preceding phrasal constituent. Bögel et al. (2008) therefore argue against treating *ezafe* as an affix and in favor of treating it as a clitic. In order to model the conflicting syntactic and prosodic properties of *ezafe*, they build on Butt and King (1998) and integrate a prosodic projection into the analysis.

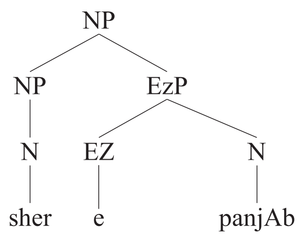
The prosodic projection proposed by Butt and King (1998) and used in the analysis of Urdu *ezafe* by Bögel et al. (2008) follows the standard architecture of LFG, which projects other levels of representation from the c-structure. Under this view, syntax is primary and the c-structure (and f-structure) are central to analyses in other projections. This arrangement is therefore in line with approaches that assume that prosodic and syntactic units are typically aligned.

Example (9) illustrates a typical *ezafe* construction. There is a syntactic dependency between the head noun *sher* ‘lion’ and a modifier to the right of that NP, *panjAb* ‘Punjab’.[5] This dependency relation is licensed by the *ezafe* (=*e* in (9)) even though the *ezafe* is prosodically attached to the head of the construction. Note that the usual pattern of headedness in Urdu (as in Persian) is head-final.

(9) sher=e panjAb
lion=Ez Punjab
‘a/the lion of Punjab’

Bögel et al. (2008) analyze this construction as follows. Syntactically, the *ezafe* is part of the modifying construction. It licenses the modifier *panjAb* and is therefore part of the same constituent. The category EzP is assigned to that constituent to model the idea that *ezafe* is a head that licenses a complement, namely the *panjAb* in our example.

(10) a. C-Structure



b. F-Structure

$$\begin{bmatrix} \textsc{pred} & \text{‘sher’} \\ \textsc{mod} & \left\{ \begin{bmatrix} \textsc{pred} & \text{‘panjAb’} \\ \textsc{gend} & \text{masc, } \textsc{num} \text{ sg, } \textsc{pers} \text{ 3} \end{bmatrix} \right\} \\ \textsc{check} & \begin{bmatrix} \textsc{ezafe} & + \end{bmatrix} \\ \textsc{gend} & \text{masc, } \textsc{num} \text{ sg, } \textsc{pers} \text{ 3} \end{bmatrix}$$

At the c-structure level shown in (10a), the *ezafe* is inserted as a terminal node and is thus analyzed as a syntactic word in its own right.[6] It combines with its complement to form the modifying constituent for the head noun *sher*. This modification relationship is expressed within the f-structure, which models the functional information and dependencies. In (10b), *sher* is the head of the phrase and *panjAb* functions as its modifier (MOD).

[5]The analysis of Urdu *ezafe* is part of on-going work on building a computational grammar of Urdu within the ParGram project (Butt et al., 1999; Butt and King, 2007). The representations below reflect the output of the implemented grammar, and the Urdu examples are provided in the ASCII transliteration scheme used by the grammar.

[6]This analysis contrasts with proposals that treat items like *ezafe* as phrasal affixes (Anderson, 2005) that do not appear as separate syntactic elements but are instead morphologically incorporated into their hosts. This is because they share some morphological properties with inflectional affixes (Zwicky, 1987; Samvelian, 2007; Miller, 1992). Most proponents of this approach have worked within Head-Driven Phrase Structure Grammar, but this type of analysis is also compatible with GB/Minimalism in that clitics can be seen as functional items that are placed high in the tree (e.g. within IP) and can thus be thought of as postlexical inflectional items (e.g. van der Leeuw (1997)). The idea of treating a subclass of clitics in the morphology has also appeared in earlier LFG-oriented proposals (e.g. Sadler and Spencer, 2000; Luís and Otoguro, 2005). On our new account this mixture of prosodic/morphological/syntactic properties follows more straightforwardly from the interaction of separate components.

However, this syntactic analysis is at odds with the fact that the *ezafe* is prosodically incorporated into the head noun to its left. That is, we have a clear mismatch between syntactic and prosodic constituency:

(11)	Syntactic Phrasing:	[[sher]	[e	panjAb]]
	Prosodic Phrasing:	((sher	e)	panjAb)

Bögel et al. (2008) make use of the Butt and King (1998) p(rosodic)-structure to address this problem and arrange for the p-structure in (12) to be assigned to this example.

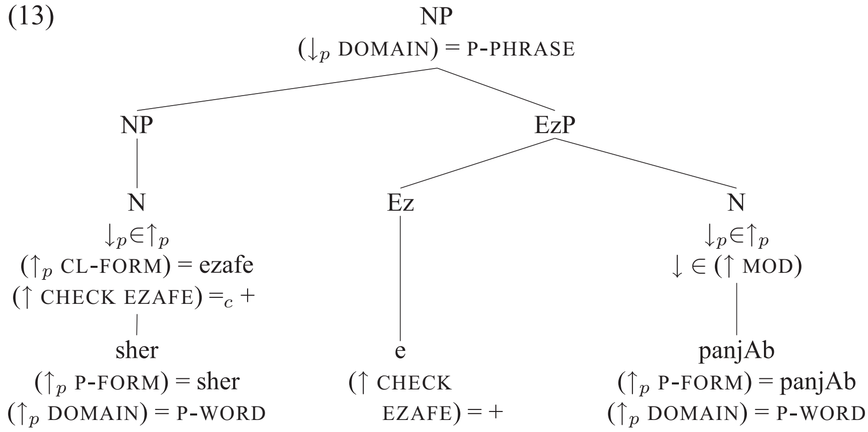
(12) P-Structure for *sher e panjAb*

$$\left[\begin{array}{l}\left\{\begin{array}{l}\left[\text{DOMAIN P-WORD, P-FORM sher, CL-FORM ezafe}\right]\\ \left[\text{DOMAIN P-WORD, P-FORM panjAb}\right]\end{array}\right\}\\ \text{DOMAIN P-PHRASE}\end{array}\right]$$

The outer unit of this p-structure corresponds to the top-level NP in (10a) and is marked as a prosodic phrase. The fact that this phrase covers the two prosodic words is formalized by the set containing the individual p-structures corresponding to those words. Crucial to this analysis, *ezafe* is not encoded as an independent prosodic word in the p-structure. It appears instead as a CL(itic)-FORM in the word-level p-structure corresponding to its prosodic host.

Bögel et al. associate the p-structure in (12) with the c- and f-structures in (10) by adding to the conventional LFG syntactic rules a set of constraints that describe the p-structure alongside the constraints that characterize the f-structure. The grammar constraints most relevant to this discussion are shown as annotations on the nodes in (13), a decorated version of the c-structure (10a). In these constraints the designators $\uparrow$ and $\downarrow$ denote the f-structures corresponding to mother and daughter nodes, as usual, and $\uparrow_p$ and $\downarrow_p$ denote the p-structure units projected from those nodes. Unless otherwise specified and as is conventional for LFG, the f- and p-structures corresponding to a daughter node are assumed by default to be the same as the structures corresponding to its mother. Thus all of the p-structure constraints under the left NP must hold of a single structure; that word-level structure is a member of the set in the top-level p-structure by virtue of the $\downarrow_p \in \uparrow_p$ assertion on the N. Similarly, the collection of constraints on the N under the EzP node define its properties as another prosodic-word component of the larger prosodic phrase. The f-structure in (10b) satisfies all of the functional constraints from the left NP (by the default convention) and also includes the MOD structure by virtue of the constraints below the *panjAb* noun.

(13)



What stands out in this representation is that the constraint that adds the CL-FORM feature to the host p-structure is attached to the N under the left NP and not to any of the nodes on the clitic side. This is because the p-structure of the leftward host is not accessible by ordinary co-description from any of the nodes on the right. Constraints on the right can make reference to the top-level p-structure, which does contain the *sher* structure as a set-element, but there is no co-descriptive designator by which that particular element can be picked out from other elements that might also belong to that prosodic phrase. The solution shown in this tree is to assign the CL-FORM on the left side, where the host p-structure is directly available. This requires an alternative expansion of the general NP rule that attaches the CL-FORM p-structure constraint to the N head but only when the NP is part of an *ezafe* construction. Relying on the fact that the top-level f-structure (unlike the p-structure) is accessible on both branches, Bögel et al. impose this restriction by having *ezafe* register its presence by adding a special CHECK feature to the top f-structure, and then testing for that feature via the $(\uparrow$ CHECK EZAFE$) =_c +$ constraint where the CL-FORM is assigned.[7]

Although this co-descriptive arrangement of c-, f- and p-structures does model the properties of the Urdu *ezafe*, this account is unsatisfactory in several ways. The grammar under this analysis does not express the pretheoretic intuition that a clitic operates on its host and instead makes the host anticipate that it might have an attached clitic. This leads to a complicated and carefully orchestrated distribution of prosodic and syntactic constraints across both lexical entries and syntactic rules. And these constraints are special to this particular construction and make no further predictions about the interaction of prosodic structure with the rest of the grammar. Other phenomena, for example case clitics or focus clitics (Butt and King, 2004), would have to be modeled individually, on a case by case

[7]CHECK features are used by convention in the ParGram grammars to encode information that is needed to ensure syntactic well-formedness but is not theoretically interesting and is not relevant to other modules or domains of application (e.g., semantic interpretation or machine translation). Bögel et al. extend the CHECK convention to handle the cross-module interaction of syntax and prosody.

basis.[8] In contrast, the architecture we develop in the next section allows prosodic generalizations to be stated independently of syntactic ones and does not require otherwise unmotivated structures or bookkeeping features to correctly model the interactions of these linguistic subsystems.

3 A Pipeline Architecture

A growing body of evidence, some of which we have summarized, calls into question the hypothesis of strong alignment between prosody and syntax and thus also the LFG co-description account of the relationship between these components. The challenge is to define an architecture of components that allows a close linkage between prosodic and syntactic phenomena in some situations but still allows for independent operation in cases of misalignment.

We suggest that this challenge can be met with a pipeline arrangement of independent components that interact through a very simple channel of communication. Our proposal depends on pre-existing aspects of the LFG syntactic formalism, including the capability of expressing optimality-theoretic preferences to impose soft constraints on syntactic interpretations (Frank et al., 1998; Sells, 2001; Bresnan, 2000). Our architecture has the following key features:

(14) a. An independent prosodic component interprets various phonological properties to determine the boundaries of prosodic phrases.

b. Prosodic boundaries are made visible to the syntax as distinct symbols in the terminal string of the syntactic constituent structure.

c. Prosodic boundary symbols augment but do not disrupt syntactic patterns.

d. The syntactic component obeys a Principle of Prosodic Preference: syntactic structures with constituent boundaries that do not coincide with prosodic boundaries are dispreferred.

[8]Asudeh (2009) proposes an elaboration of the LFG formalism that allows constraints associated with lexical nodes to make direct reference to the structures that correspond to preceding or following lexical nodes. He aims to account for the restrictions on Complementizer-Adjacent Extraction (e.g. *that*-trace and fixed-subject constraints), but his technique might offer a simpler and more intuitive account of clitic prosodic attachment within a configuration of co-described representations. Asudeh builds on a suggestion made originally by Kaplan (1987, 1989) to formalize the mapping between phonological tokens and the lexical nodes of the c-structure in terms of another projection function within the overall Correspondence Architecture of LFG. He observes that this function, denoted as π, is one-to-one, and its inverse is therefore a function that maps from a lexical node to the corresponding phonological token. This can be composed with functions that take phonological tokens into the tokens that precede or follow them. Asudeh defines a new designator $\succ = \phi(M(\pi(Next(\pi^{-1}(*)))))$ to designate the f-structure of a following lexical node. Similarly, we can define $\prec_p$ as a designator for the p-structure corresponding to the lexical node of a preceding phonological token. Given this machinery, we can replace the CHECK assignment on the EZAFE clitic with the constraint ($\prec_p$ CL-FORM) = ezafe and remove the CL-FORM and CHECK annotations from the host noun.

This formalization avoids some of the unintuitive aspects of the Bögel et al. (2008) account, but it still requires a case-by-case distribution of constraints.

In this architecture the boundary-annotated output of the separate prosodic component becomes the input to the syntactic component. The input to the syntax for our Urdu *ezafe* example is the prosodically-bracketed string in (11) instead of the simple word-string. The syntactic component interprets its rules as allowing prosodic brackets to be freely intermixed among the other syntactically specified terminals (property (14c)), and the resulting syntactic structures have nodes and branches that cover the prosodic brackets in addition to the required syntactic formatives. This means that misaligned prosodic brackets will not interfere with the usual syntactic analysis. Thus rule (15a) is interpreted as (16a) and (15b) is interpreted as (16b) when they apply to our *ezafe* example (RB and LB are the lexical categories of the right/left prosodic brackets, the terminal parentheses).

(15) a. EzP → EZ N
$\downarrow \in (\uparrow \text{MOD})$

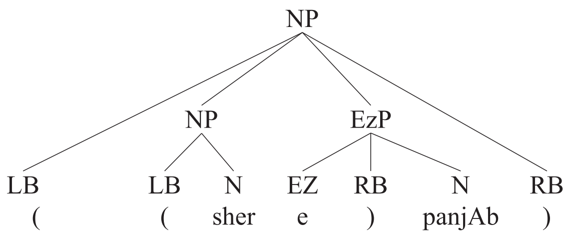
b. NP → N

(16) a. EzP → EZ RB N
$\downarrow \in (\uparrow \text{MOD})$

b. NP → LB N

This architecture allows a drastic simplification of the rules needed to describe an *ezafe* construction. In contrast to the annotations needed for the previous solution, the rules in (15) do not encode any information about prosodic properties and do not involve any CHECK-features to ensure that an *ezafe* clitic appears to the right of the head noun. With the extended rule interpretation in (16), the tree in (17) is the resulting c-structure.

(17)



The outer prosodic brackets are aligned with the syntactic constituents but the internal ones are not. The input is still accepted by the grammar and is assigned the f-structure in (10b). The traditional syntactic c-structure can be seen as a projection of (17) formed by systematically deleting prosodic nodes and branches.

Consider (18) as another illustration of misalignment in our pipeline. Again we assume that the prosodic component introduces phonologically-determined boundaries into the syntactic input string. We have added square brackets on top of the prosodically-bracketed input to indicate the syntactic constituents — the syntactic analysis goes through despite the confusion of prosodic boundaries.

(18)	Phonological String	*Drink a pint a milk a day*
	PROSODIC PARSER	|
	Prosodically Bracketed	(Drink a) (pint a) (milk a) (day)
	SYNTACTIC PARSER	|
	Syntactic Analysis	[[(Drink] [[a) (pint] [of) (milk]] [a) (day)]]

We have seen how properties (14a-c) allow for arbitrary misalignments of syntactic and prosodic boundaries. The Principle of Prosodic Preference (14d) completes the architecture by introducing a soft dependency between prosody and syntax: among a competing set of syntactic structures, those with the fewest number of misaligned brackets will be selected as the correct analyses. This reflects Selkirk's original intuition of close alignment, at least in certain situations. As an immediate consequence, it also captures the fact that prosodic information can have the effect of disambiguating between several possible parses. The phrase *old men and women* is syntactically ambiguous in the absence of prosodic phrasing, as indicated by the following syntactic structures:

(19) Syntactic constituents

a. [[old men] and [women]]

b. [old [men and women]]

But suppose that the syntax is instead given a prosodically bracketed string, for example, the one in (20a):

(20) Prosodic phrasing

a. (old men) (and women)

b. (old) (men and women)

With this prosodic phrasing the analysis in (19b) is dispreferred by virtue of the bracket configurations shown in (21), and the compatible analysis in (19a) is selected. The asterisks mark the prosodic brackets that are unaligned with syntactic phrases and are therefore dispreferred. The alternative phrasing in (20b) will select the analysis (19b).

(21) a. [[(old men)] *(and [women])]

b. [(old [men*) *(and women])]

To summarize, this architecture for the interface between prosody and syntax allows a proper analysis of the Urdu *ezafe* clitic, the systematic misalignment between prosody and syntax in *Drink a pint a . . .* , the Dutch definite determiner clitic, and, as far as we can determine, all other clitic phenomena and other instances of misalignment. We allow misalignments as a matter of course, as suggested by one school of thought on these matters, but we also incorporate a preference for more aligned analyses, in accord with the second and more conventional school of thought. Prosody operates as a separate component that communicates with syntax through the narrow channel of a prosodically annotated string. The syntactic interactions are governed by an extended interpretation of ordinary c-structure rules and an optimality-theoretic Principle of Prosodic Preference. Unlike previous LFG proposals, we do not incorporate a co-described prosodic projection and so avoid the detailed specifications that define its properties.

4 Implementation by Metarule Expansion

Our proposed architecture assigns an extended interpretation to the ordinary rules of a conventional LFG grammar. In (16) we showed that the effect of this extended interpretation for some particular rules is equivalent to including in the grammar some additional rules that are systematically related to the originals. We observe now that this is generally the case: the behavior of every syntactic rule according to our proposed architecture can be modeled by a finite expansion to a set of rules that could have been written in standard, pre-existing notations. In other words, the architectural principles in (14) can be implemented as metagrammatical operations that systematically transform the rules of a conventional grammar. As a consequence, we know that this architecture implies no changes to the mathematical and computational properties of the syntactic component.

A conventional LFG grammar contains a set of c-structure rules of the form

(22) CAT → RHS

where CAT is a nonterminal category and the right-hand side RHS denotes a regular language over categories annotated with functional (or other co-describing) constraints. To implement the architectural specifications, we replace each such rule with another rule of the form

(23) CAT → (LB) RHS / $[\underset{\text{Disprefer}}{\text{LB}|\text{RB}}]$ (RB)

The prosodic brackets and their lexical categories (LB and RB) belong to the terminal and nonterminal vocabularies of the enlarged grammar, in accordance with (14b). The right-side of the original rule is replaced by a rule expansion which allows for the parsing of prosodic brackets. The categories of the original right-hand side can be optionally preceded by a left prosodic bracket (as indicated by the parentheses) and optionally followed by a right prosodic bracket. In addition, the expansion will match a daughter sequence that would match the RHS regular expression if all occurrences of either LB or RB in that sequence are ignored (the | indicates a disjunction). The / is a notation for the "Ignore" operator first introduced by Kaplan and Kay (1994); it is included in the Xerox finite-state machine calculus (Beesley and Karttunen, 2003) and in the c-structure notation of the XLE system (Crouch et al., 2009). The effect of this use of the Ignore operator is to implement property (14c) of the architecture: it ensures that occurrences of prosodic brackets cannot disrupt otherwise valid phrase-structure expansions.

The "Disprefer" annotation implements the Principle of Prosodic Preference (14d). Whenever a prosodic bracket is ignored in the middle of the RHS, the structure is assigned a dispreference optimality mark. The effect of this is to determine a ranking over possible syntactic analyses, as described by Frank et al. (1998). The only brackets that are not dispreferred are those that match the optional LB and RB categories, the ones that appear on the edges of constituents. Replacement rules produced in this way by metagrammatical expansion thus provide dispreferences only for misaligned prosodic brackets, as required.

By way of illustration, the example in (24b) shows what results from the metarule expansion of the simple rule (24a).

```
(24)  a. VP →  V        NP
                        (↑ OBJ)=↓
      b. VP →
         (LB)  [LB|RB]*   V   [LB|RB]*       NP        [LB|RB]*   (RB)
               Disprefer      Disprefer   (↑ OBJ)=↓   Disprefer
```

The Kleene-star operators derive from the Ignore specification. They allow for misaligned prosodic brackets to appear in any position as well as for the possibility of no misalignments.

The metagrammatical implementation of our architecture can be instantiated quite directly within the XLE computational system (Crouch et al., 2009). XLE includes a metarule expansion facility whose purpose is to express generalizations over all syntactic rules or over particular subsets of them. This facility is invoked by defining a "metarule macro". The input to a metarule macro is the category and right-hand side of an existing rule, and the output is the replacement rule for that input.[9] The macro definition in XLE notation in (25) is equivalent to the metagrammatical expansion in (23).

```
(25)  METARULEMACRO(CAT, RHS) =
      (LB)  RHS / {LB|RB}:@(DISPREFER)  (RB)
```

The @(DISPREFER) annotation is an invocation of an XLE template that can be defined to add a prosodic dispreference to the collection of optimality marks associated with the c-structure.

[9] In the XLE implementation, METARULEMACRO takes three arguments: the CATegory, the BASE CATegory, and the RHS. For our purposes, the distinction between the category and the base category is unimportant.

5 The Prosodic Component: Some Speculations

Our architecture postulates an independent prosodic component that recognizes prosodic phrases and marks their boundaries in the input string to the syntax. On this view, the internal properties of the prosodic component are not accessible to syntax and are not constrained by syntactic requirements, and indeed are not especially relevant to the overall architectural conception. Still, it is worthwhile to consider how the prosodic component might operate.

The prosodic component must embody knowledge about rhythmic structure of the language (trochaic/iambic), it must be able to parse tones (e.g. high/low), and it must be sensitive to part of speech information, at least enough to differentiate function vs. content words. There have been some suggestions in the literature that recursive rules may be needed for prosodic analysis (Booij, 1995, 1996; Peperkamp, 1997; Vigário, 1999, 2003). However, the notion of recursivity within phonology seems to be confined to the level of the prosodic word and mainly seems to concern clitic phenomena across languages. More recent work seems to be distancing itself from the notion of recursivity in phonology. Vogel (2009) argues that recursive power is not necessary for clitic phenomena, and Kabak and Revithiadou (2009) attribute its appearance at the prosodic level to the interaction with morphosyntax. That is, recursivity within morphosyntax is reflected within prosody but is not inherent to prosodic structures (see also Selkirk's (1984) *Strict Layer Hypothesis*, which legislates against recursion). O'Connor (2004) also points out that center-embedding recursion is not needed for prosodic structure.

If prosodic rules lack center-embedding recursion and have a bounded number of levels (e.g. prosodic words, prosodic phrases, intonational phrases), then the prosodic component as a whole defines a regular relation between its inputs and outputs, a relation that can be implemented by a finite-state transducer. Thus O'Connor (2005b) proposes to model prosodic information via a series of rewrite rules that apply to the representation of intonation in the AM/ToBI annotation scheme (Pierrehumbert, 1980; Goldsmith, 1976, among others). His proposal allows only bounded reapplication of these tune structure rules and so his system can describe only regular relations.

These observations have some interesting and important consequences. If the prosodic component defines only regular relations, we can characterize it using notational devices whose mathematical and computational properties are very well understood (Kaplan and Kay, 1994; Beesley and Karttunen, 2003). We also know that the LFG languages are closed under composition with regular relations, so that the formal power of the combination of components in our pipeline architecture is no greater than the formal power of the syntactic component by itself.

As another consequence, we can immediately create and experiment with a concrete implementation of our architecture. In the XLE system the terminal string of the syntactic tree is constructed by applying a pipeline sequence of finite state transductions to an original input string. Typically the input is a string of ordinary text, and the transducers perform standard transformations such as tokenization and morphological analysis. We can reconfigure the system so that its input is a string annotated with tonal information and other prosodically-relevant features. Then the initial step in the cascade of transformations can be carried out by a transducer that introduces prosodic brackets that are consistent with the prosodic annotations. We are now experimenting with a first version of this type of transducer, constructed using the tools of the XFST finite-state calculus (Beesley and Karttunen, 2003).

It is a strong and useful hypothesis that the independent prosodic component is so limited in its computational power, but it is not a theoretical necessity of our proposed architecture.

6 Conclusion

We have proposed an LFG-oriented pipeline architecture that allows for misalignments between prosody and syntax in a natural manner but also still incorporates a preference for an alignment of prosodic and syntactic phrases. We postulate that an independent prosodic component delivers prosodically-bracketed strings as the input to the syntactic component and that syntax can ignore these brackets with some degree of dispreference if they are incompatible with proper syntactic analyses. This architecture provides explicit accounts of the syntactic and prosodic properties of clitics using simpler rules and representations than previous approaches have required. We have also shown how this architecture can be implemented by means of metagrammatical expansions, both conceptually and computationally, so that it adds no new formal power to the basic LFG framework. This architecture addresses the challenges coming from the two traditional schools of thought concerning the alignment of prosody and syntax by allowing both for the primacy of syntax and for rampant mismatches between syntactic and prosodic structure.

References

Anderson, Stephen. 2005. *Aspects of the Theory of Clitics*. Oxford University Press.

Asudeh, Ash. 2009. Adjacency and Locality: A Constraint-based Analysis of Complementizer-Adjacent Extraction. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG09 Conference*, CSLI On-line Publications.

Beesley, Kenneth and Karttunen, Lauri. 2003. *Finite State Morphology*. CSLI Publications.

Berendsen, Egon. 1986. *The Phonology of Cliticization*. Foris.

Bögel, Tina, Butt, Miriam and Sulger, Sebastian. 2008. Urdu Ezafe and the Morphology-Syntax Interface. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG08 Conference*, CSLI On-line Publications.

Booij, Geert. 1995. *The Phonology of Dutch*. Clarendon Press.

Booij, Geert. 1996. Cliticization as Prosodic Integration: The Case of Dutch. *The Linguistic Review* 13, 219–242.

Bresnan, Joan. 2000. Optimal Syntax. In Joost Dekkers, Frank van der Leeuw and Jeroen van de Weijer (eds.), *Optimality Theory: Phonology, Syntax, and Acquisition*, pages 334–385, Oxford University Press.

Butt, Miriam and King, Tracy Holloway. 1998. Interfacing Phonology with LFG. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG98 Conference*, CSLI On-line Publications.

Butt, Miriam and King, Tracy Holloway. 2004. The Status of Case. In Veneeta Dayal and Anoop Mahajan (eds.), *Clause Structure in South Asian Languages*, Kluwer.

Butt, Miriam and King, Tracy Holloway. 2007. Urdu in a Parallel Grammar Development Environment. *Language Resources and Evaluation* 41, 191–207.

Butt, Miriam, King, Tracy Holloway, Niño, María-Eugenia and Segond, Frédérique. 1999. *A Grammar Writer's Cookbook*. CSLI Publications.

Carlos Gussenhoven, C. 1989. Cliticization in Dutch as Phonological Word Formation, University of Nijmegen.

Cinque, Guglielmo. 1993. A Null Theory of Phrase and Compound Stress. *Linguistic Inquiry* 24, 239–267.

Crouch, Dick, Dalrymple, Mary, Kaplan, Ronald M., King, Tracy, Maxwell, John and Newman, Paula. 2009. *XLE Documentation*. On-line documentation, Palo Alto Research Center (PARC).

Dalrymple, Mary, Kaplan, Ronald M., Maxwell III, John T. and Zaenen, Annie (eds.). 1995. *Formal Issues in Lexical-Functional Grammar*. CSLI Publications.

Frank, Anette, King, Tracy Holloway, Kuhn, Jonas and Maxwell III, John T. 1998. Optimality Theory Style Constraint Ranking in Large-scale LFG Grammars. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG98 Conference*, CSLI On-line Publications.

Goldsmith, John. 1976. *Autosegmental Phonology*. Ph. D. thesis, Massachusetts Institute of Technology.

Gussenhoven, Carlos. 1986. Over de Fonologie van Nederlandse Clitica. *Spektator* 15, 18–200.

Halpern, Aaron. 1995. *On the Placement and Morphology of Clitics*. CSLI Publications.

Halpern, Aaron and Zwicky, Arnold (eds.). 1996. *Approaching Second: Second Position Clitics and Related Phenomena*. CSLI Publications.

Hayes, Bruce and Lahiri, Aditi. 1991. Bengali Intonational Phonology. *Natural Language and Linguistic Theory* 9, 47–96.

Kabak, Bariş and Revithiadou, Anthi. 2009. An Interface Approach to Prosodic Word Recursion. In Janet Grijzenhout and Bariş Kabak (eds.), *Phonological Domains: Universals and Deviations*, pages 105–133, Mouton de Gruyter.

Kaplan, Ronald M. 1987. Three Seductions of Computational Psycholinguistics. In Peter Whitelock, Mary McGee Wood, Harold L. Somers, Rod Johnson and Paul Bennett (eds.), *Linguistic Theory and Computer Applications*, pages 149–181, Academic Press, reprinted in Dalrymple et al. (1995:339–367).

Kaplan, Ronald M. 1989. The Formal Architecture of Lexical-Functional Grammar. In Chu-Ren Huang and Keh-Jiann Chen (eds.), *Proceedings of ROCLING II*, pages 3–18, reprinted in Dalrymple et al. (1995:7–27).

Kaplan, Ronald M. and Kay, Martin. 1994. Regular Models of Phonological Rule Systems. *Computational Linguistics* 20, 331–478.

Lahiri, Aditi and Fitzpatrick-Cole, Jennifer. 1999. Emphatic Clitics and Focus Intonation in Bengali. In René Kager and Wim Zonneveld (eds.), *Phrasal Phonology*, Nijmegen University Press.

Lahiri, Aditi, Jongman, Allard and Sereno, Joan A. 1990. The Pronomincal Clitic [der] in Dutch: A Theoretical and Experimental Approach. *Yearbook of Morphology* 3, 115–127.

Lahiri, Aditi and Plank, Frans. 2009. What Linguistics Universals can be True of. In Sergio Scalise, Elisabetta Magni and Antonietta Bisetto (eds.), *Universals of Language Today*, pages 31–58, Springer.

Luís, Ana and Otoguro, Ryo. 2005. Morphological and Syntactic Well-formedness: The Case of European Portuguese Clitics. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG05 Conference*, CSLI On-line Publications.

Miller, P. 1992. *Clitics and Constituents in Phrase Structure Grammar*. Garland.

Mycock, Louise. 2006. *The Typology of Constituent Questions: A Lexical-Functional Grammar Analysis of wh-questions*. Ph. D. thesis, University of Manchester.

Nespor, Marina and Vogel, Irene. 1986. *Prosodic Phonology*. Foris.

O'Connor, Rob. 2004. *Information Structure in Lexical-Functional Grammar: The Discourse-Prosody Correspondence in English and Serbo-Croation*. Ph. D. thesis, University of Manchester.

O'Connor, Rob. 2005a. Clitics in LFG: Prosodic Structure and Phrasal Affixation. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG05 Conference*, CSLI On-line Publications.

O'Connor, Rob. 2005b. Information Structure in Lexical-Functional Grammar: The Discourse-Prosody Correspondence in English and Serbo-Croation. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG05 Conference*, CSLI On-line Publications.

Peperkamp, Sharon. 1997. *Prosodic Words*. Ph. D. thesis, University of Amsterdam.

Pierrehumbert, Janet. 1980. *The Phonology and Phonetics of English Intonation*. Ph. D. thesis, Massachusetts Institute of Technology.

Plank, Frans. 2005. ((FRISCH ge) (WAGT ist)) ((HALB ge) (WON nen)), unpublished manuscript, Universität Konstanz.

Sadler, Louisa and Spencer, Andrew. 2000. Workshop on Morphology and LFG. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG00 Conference*, CSLI On-line Publications.

Samvelian, Pollet. 2007. A (Phrasal) Affix Analysis of the Persian Ezafe. *Journal of Linguistics* 43, 605–645.

Selkirk, Elisabeth. 1981. On Prosodic Structure and its Relation to Syntactic structure. In Thorstein Fretheim (ed.), *Nordic Prosody II*, pages 111–140, Tapir.

Selkirk, Elisabeth. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. MIT Press.

Selkirk, Elisabeth. 1986. On Derived Domains in Sentence Phonology. *Phonology Yearbook* 3, 371–405.

Selkirk, Elisabeth. 1995. The prosodic structure of function words. In Jill Beckman, Laura Walsh Dickey and Suzanne Urbanczyk (eds.), *Papers in Optimality Theory*, pages 439–70, GLSA Publications, also in Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition, ed. James L. Morgan and Katherine Demuth, 187-214: Lawrence Erlbaum Associates.

Selkirk, Elisabeth. 2001. The syntax-phonology interface. In N.J. Smelser and Paul B. Baltes (eds.), *International Encyclopedia of the Social and Behavioral Sciences*, pages 15407–15412, Pergamon.

Sells, Peter (ed.). 2001. *Formal and Empirical Issues in Optimality Theoretic Syntax*. CSLI Publications.

Truckenbrodt, Hubert. 1999. A Null Theory of Phrase and Compound Stress. *Linguistic Inquiry* 30, 219–255.

van der Leeuw, Frank. 1997. *Clitics: Prosodic Studies*. Holland Academic Graphics.

Vigário, Marina. 1999. On the Prosodic Status of Stress Function Words in European Portuguese. In Tracy Alan Hall and Ursula Kleinhenz (eds.), *Studies on the Phonological Word*, pages 253–293, John Benjamins.

Vigário, Marina. 2003. *The Prosodic Word in European Portuegese*. Mouton de Gruyter.

Vogel, Irene. 2009. The Status of the Clitic Group. In Janet Grijzenhout and Bariş Kabak (eds.), *Phonological Domains: Universals and Deviations*, Mouton de Gruyter.

Wheeldon, Linda and Lahiri, Aditi. 1997. Prosodic Units in Speech Production. *Journal of Memory and Language* 37, 356–381.

Wheeldon, Linda and Lahiri, Aditi. 2002. The Minimal Unit of Phonological Encoding: Prosodic or Lexical Word. *Cognition* 85, B31–B41.

Zwicky, Arnold. 1987. Suppressing the Z's. *Journal of Linguistics* 23(1), 133–148.