

Chapter 17

Prosody and its interfaces

Tina Bögel

University of Konstanz

LFG has always had a strong focus on syntax and semantics, but the last two decades have seen significant progress with regard to the integration of p(honological)-structure into LFG. This chapter first briefly introduces important concepts for the analysis of prosody and gives an overview of widely adopted approaches to the syntax–prosody interface. The second part surveys the different proposals for the integration of p-structure and its interfaces into LFG, with a particular focus on the architectural assumptions behind each approach and the resulting implications for the architecture of grammar.

1 Introduction

LFG has always had a strong focus on syntax and semantics. However, with the realisation that prosodic information can significantly contribute to linguistic analyses and is often crucial for the correct interpretation of meaning (e.g., in form of prosodic disambiguation of syntactically ambiguous structures or for the correct interpretation of information structure), the last two decades have seen significant progress with regard to the integration of prosodic structure into LFG.

LFG assumes that different aspects of grammar (i.e., syntax, semantics, etc.) are represented by unique modules (also called ‘projections’), each guided by its own principles and constraints, and with representations well-suited to their unique functions (cf. Dalrymple 2001; Sadock 1991, see also Belyaev forthcoming(b) [this volume]). The syntactic component, for example, is represented by c(onstituent)- and f(unctional)-structure and is concerned with constituency (via phrase structure rules) and the encoding of grammatical functions and morphosyntactic features, while phonology (including prosody) is represented by p(honological)-

structure and is concerned with phonological and prosodic properties like prosodic phrasing, rhythmic constraints, and intonation.¹

Communication between the different modules is handled by LFG's correspondence architecture, which allows for relevant information to be made available at the respective interfaces. The establishment of these interfaces necessarily presumes a specific grammar architecture; that is, it presupposes an explicit positioning of modules with respect to each other. Discussing the architectural assumptions made in each p-structure proposal is thus essential for the understanding of the (in parts fundamental) differences in the representation of prosody and the communication at the interfaces.

This chapter provides an overview of the different approaches to prosody and its interfaces in LFG, and places these with respect to proposals made in the wider literature. It furthermore discusses the architectural assumptions made in each proposal and offers insights into a more general view of grammar. The chapter is structured as follows: Section 2 provides a general introduction into two major aspects of prosody (phrasing and intonation) and discusses current approaches to prosody and its interfaces in the wider literature. A discussion of the LFG grammar architecture and the placement of the phonological module (including prosody) is provided in Section 3. This section also establishes a fundamental difference between the proposals with respect to how grammar is viewed in general. Section 4 provides a chronological overview of the different approaches to prosody and its interfaces in LFG, in particular with respect to the architectural assumptions made in each proposal. Section 5 concludes the chapter.

2 Prosody and its interfaces

The LFG approaches to prosody discussed in Section 4 draw on a number of notions and theories established in the wider literature. This section first gives an overview on the general features that are particularly relevant with respect to the analysis of prosody at the interfaces and then describes the major approaches to the interface between syntax and prosody.

¹In most of the approaches discussed in Section 4 the 'p' in p-structure represents p(rosodic)-structure, as prosodic features usually contain relevant information for analyses at the interfaces to syntax, semantics, and information structure. However, prosody is only one part of the larger field of phonology and some phenomena that are not part of prosody (e.g., postlexical sandhi phenomena) can be closely interlaced with prosody in that they can indicate a specific prosodic domain. This chapter will thus use the term p(honological)-structure, of which prosody is part, but which does not, per se, restrict p-structure to represent prosody alone.

2.1 What is prosody?

Prosody is a term used to describe suprasegmental phonology. It goes beyond the phonemic level of segmental phonology and is concerned with larger units in spoken language, including prosodic grouping, intonation and/or tones, rhythm, and stress patterns. Prosody can be used to express a number of properties and functions, among these clause type, clause structure, semantic scope, concepts of information structure such as topic and focus, but also speaker emotion, irony, or sarcasm. A detailed description of prosody goes far beyond the aim of this chapter and this section will only focus on some basic notions of prosody deemed fundamental for the current state of the art in LFG, namely prosodic phrasing, intonation, and the relationship between prosody and other modules of grammar.

Traditionally, it is assumed that spoken language is grouped into hierarchically structured **prosodic domains** (e.g., Selkirk 1978; Nespor & Vogel 1986; Hayes 1989). Example (1) shows the most widely used proposal for the prosodic hierarchy originally made in Selkirk (1978) (building on an earlier proposal by McCawley 1968; see also Frota 2012 for different suggestions).

(1) The Prosodic Hierarchy (Selkirk 1978)

Prosodic hierarchy

U	utterance (Ut)
ι	intonational phrase (IntP)
φ	phonological phrase (PhP)
ω	prosodic word (PW)
Σ	foot (Ft)
σ	syllable (Syll)

In addition, the constraints in (2) are assumed to apply to the prosodic hierarchy.²

(2) **Constraints on Prosodic Domination (Selkirk 1995: ex. 4)**
 (where C^n = some prosodic category)

²These constraints have been challenged and are now mostly considered to be ‘soft’ constraints, see, e.g., Bennett & Elfner (2019).

- (i) *Layeredness*: No C^i dominates a C^j , $j > i$,
e.g. “No syllable dominates a foot.”
- (ii) *Headedness*: Any C^i must dominate a C^{i-1} (except if C^i = syllable),
e.g. “A prosodic word must dominate a foot.”
- (iii) *Exhaustivity*: No C^i immediately dominates a constituent C^j , $j < i-1$,
e.g. “No prosodic word immediately dominates a syllable.”
- (iv) *Nonrecursivity*: No C^i dominates C^j , $j=i$,
e.g. “No foot dominates a foot.”

The identification of a prosodic unit is based on various types of evidence and can vary greatly across languages. Among these types of evidence are sandhi processes (e.g., linking and intrusive /r/ in English (Wells 1970)), tonal events (e.g., Beckman & Pierrehumbert 1986; Pierrehumbert & Beckman 1988), and rhythmic patterns (e.g., Liberman 1975; Nespor & Vogel 1989). A phonological phrase in English, for example, is assumed to be intonationally represented by a pitch accent and a phrase accent, and to show phrase-final lengthening, where the last syllable is significantly longer compared to the other syllables of the phonological phrase (Lehiste et al. 1976; Frota 2012).

Tonal events like accents and boundary tones can contribute significantly to the meaning of a clause. These events are often described in terms of High and Low tones and tone combinations following the ToBI annotation conventions.³ The first set of conventions was developed for American English in 1992 (Silverman et al. 1992); others have followed with specific adaptations to other languages (e.g., German GToBI (Grice & Baumann 2002)). The ToBI conventions distinguish between three tonal events:

- Pitch accents (L^* and H^* , and combinations like $L+H^*$ and L^*+H) are usually found on the words that are most important for an interpretation. In a neutrally pronounced sentence like *Amra went to the playground to meet her friends*, ‘Amra’, ‘playground’ and ‘friends’ would usually carry pitch accents. Pitch patterns can reflect information structure (Zaenen forthcoming [this volume]): Contrastive focus in Germanic languages, for example, can be indicated by the use of an accent with a notably larger pitch span compared to the other accents of the clause (see, e.g., Féry 2020).

³The Autosegmental-Metrical/Tone and Break Indices framework (AM/ToBI) (Pierrehumbert 1980; Silverman et al. 1992; Beckman et al. 2005) is a generally adopted set of conventions to describe tonal events in the intonational contour. Break indices, which indicate the strength of a break between words, are not further discussed in this chapter.

- Boundary tones (H% and L%) are only associated with phrase edges of larger prosodic units, most often the intonational phrase boundary. They can, for example, signal the difference between a question and a statement with identical linear word order by means of rising or falling final phrase boundary tones.
- Phrase accents (H- and L-) are situated between a pitch accent and a boundary tone. They are often related to the edge of a prosodic domain below the intonational phrase, but there is some variation (see the discussion in Grice et al. 2000). They can significantly contribute to the disambiguation of syntactically ambiguous structures.

While these conventions are adopted by the vast majority of the field, proposals with a more fine-grained understanding of tonal events in combination with, for example, a distinct level of prominence (which is essential for the interpretation of focus type), have recently been developed (e.g., DIMA: Kügler et al. 2019). Whether these proposals allow for a more thorough interpretation of prosody and meaning is subject to future research.

Both areas, prosodic constituency and intonation, are deeply intertwined with each other, and are also closely associated with **segmental phonology**, in that phonological processes (e.g., resyllabification) may be constrained to a particular prosodic domain (e.g., the phonological phrase), or the quality of a vowel may change if it is associated with a pitch accent. Segmental and suprasegmental prosody both are part of lexical and postlexical phonology. Prosodic constituency and (lexical) stress are also part of a word's lexical entry, as is the knowledge about prosodically deficient clitics, while segmental phenomena frequently also occur between two words and hence are not restricted to the lexicon. As a consequence, p-structure should not only represent prosodic structure, but should rather include lexical and postlexical segmental and suprasegmental phonology (cf. fn 1).

Phonetics can be viewed as the physical translation of phonology into a concrete speech signal (and vice versa), which is reflected in the close relationship between prosodic terms like pitch, length, or loudness, and phonetic terms like fundamental frequency, duration, and intensity (see Kingston 2019 for a detailed discussion). Phonetics has not been in the focus of the proposals made in LFG, although initial approaches towards its integration have been undertaken (see Butt et al. 2020; Bögel 2020; also Section 4.5).

It is clear that prosodic structure is governed by p-structure internal principles and constraints, and that, for example, rhythm and the prosodic status of words

(prosodic words vs. prosodically deficient clitics) can determine the formation of prosodic domains. It is, however, equally assumed that prosody is influenced by syntactic structure and discourse-related aspects like the differentiation between new and given information and the expression of different focus-types (Zaenen forthcoming [this volume]). Furthermore, ‘extralinguistic’ factors such as speaking rate or frequency effects can affect p-structure.⁴ The exact influence of these and other factors on prosody is far from being fully explored. As the vast majority of research (within and outside of LFG) has focussed on the exploration of the relationship between syntax and prosody, the major approaches to this interface are briefly introduced here, before turning to the role of p-structure and the different proposals to prosody and its interfaces in LFG.

2.2 Theories of the prosody-syntax interface

The literature on how the syntactic and the prosodic modules interact can be roughly divided into two major camps: **direct reference** and **indirect reference** (see Bennett & Elfner 2019 for a detailed discussion of each approach). The direct reference approach proposes that phonological rules and groupings can directly be conditioned on syntactic relations or properties, e.g., on c-command, sister relations, or ‘head’ status, without the intervention of a separate prosodic structure (e.g., Kaisse 1985; see Elordieta 2008 for an overview). As LFG assumes a modular view of grammar and none of the LFG approaches propose the (non-modular) direct reference approach, this chapter will not further discuss this particular approach to the interface.

The other school of thought pursues the indirect reference approach, which assumes that syntactic structure is first mapped to prosodic domains as shown in (1). Phonological rules are then conditioned based on these prosodic domains (e.g., Hayes & Lahiri 1991). Prominent proposals include the **end-based** approach (Selkirk 1986; Chen 1987) which assumes that the mapping algorithm is restricted to the edges of syntactic heads and maximal projections. In the abstract example in Figure 1, each syntactic head receives a prosodic word boundary and each XP receives a phonological phrase boundary at its right edge. As all XPs align at their right edge, only one phonological phrase boundary is included.⁵ Function words (‘fw’) are excluded from the mapping algorithm.

⁴See Shattuck-Hufnagel & Turk (1996) for a thorough discussion of different constraints on prosody.

⁵Whether the right or the left edge is aligned seems to be subject to language-specific constraints.

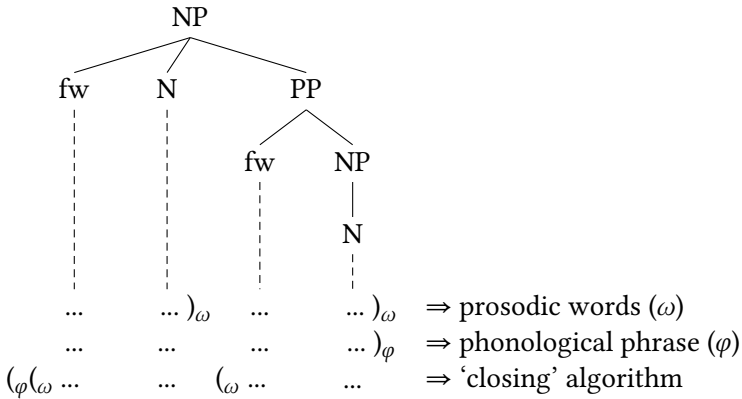


Figure 1: The end-based approach (Selkirk 1986: 387, shortened and modified)

All prosodic domains in this example are ‘closed’ by assuming an automatic insertion of a left boundary with the neighbouring right boundaries or at the edges of the whole construction. This effectively groups all function words together with their corresponding syntactic heads into prosodic words and places both prosodic words within the phonological phrase: $(\phi(\omega \text{fw N})_{\omega}(\omega \text{fw N})_{\omega})_{\phi}$.

The end-based approach has been reformulated as a generalized alignment constraint in Optimality Theory (OT; McCarthy & Prince 1993; Prince & Smolensky 2004) and is generally represented in this format, for example as ALIGN-XP-R(IGHT): ALIGN (XP, R; ϕ , R)⁶ (Selkirk 1995).

In later work, Selkirk (2009, 2011) introduced **match theory** (which has ancestors in, e.g., Ladd 1986). In contrast to the previous end-based approach, match theory assumes that both edges of a syntactic constituent are simultaneously matched to a prosodic constituent.

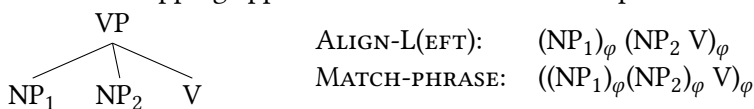
- MATCH-CLAUSE: A clause in syntactic constituent structure must be matched by a corresponding intonational phrase (ι) in prosodic constituent structure: MATCH (CLAUSE, ι)
- MATCH-PHRASE: A phrase in syntactic constituent structure must be matched by a corresponding phonological phrase (ϕ) in prosodic constituent structure: MATCH (XP, ϕ)

⁶Read as: “Align the right edge of an XP with the right edge of a phonological phrase/ ϕ ”.

- MATCH-WORD: A (lexical) word in syntactic constituent structure must be matched by a corresponding prosodic word (ω) in prosodic constituent structure: MATCH (LEX-WORD, ω) (Selkirk 2011: 439, modified)

Match theory reflects the syntactic structure in much more detail; in particular (and in contrast to the end-based approach) it predicts recursion, as syntactically nested XPs will be phrased as recursive structures in prosodic constituency. The (modified) example in (3) from Selkirk (2011) illustrates this point for a transitive verb phrase.

- (3) Different mapping approaches for a transitive verb phrase



In (3), each VP/NP receives a preceding left phonological phrase boundary in the end-based approach (ALIGN-L). The φ -boundary for NP₁ is identical with the boundary for the VP. The right boundary for NP₁ is placed ('closed') before NP₂, and the second right φ -boundary is placed after V, which does not receive any boundaries by itself. The first phonological phrase thus contains NP₁ and the second phonological phrase groups the verb together with NP₂. In contrast, the MATCH algorithm maps each XP (NP₁, NP₂, and VP) into a phonological phrase, thus creating a recursive structure.

Besides these two major schools of direct and indirect reference, a third proposal with regard to the formation of prosodic structure has been adopted into the LFG community as well (most prominently in Dalrymple & Mycock 2011; Mycock & Lowe 2013, and subsequent work), which will be called the **parallel approach** in this chapter. The main motivation for the parallel approach is the frequently observed non-isomorphism between syntactic and prosodic constituency as illustrated in example (4).

- (4) Syntactic Phrasing: [Drink [[a pint] [of milk]] [a day]]
 Phonological Phrasing: (Drink a) (pint a) (milk a) (day)
 (Lahiri & Plank 2010: 376, modified)

This frequent mismatch seemingly rules out any approaches which map syntactic constituents to prosodic domains, but suggests that prosodic structure is built up on prosodic principles alone. Based on observations of rhythmic patterns (e.g., Sweet 1904), Lahiri & Plank (2010) assume the trochaic foot (X –) to be the determining element for the creation of prosodic structure in English,

with the stressed syllable as the initial element of each prosodic chunk. Lahiri and Plank discuss this approach to prosodic phrasing with respect to a number of diachronic and synchronic examples which support the assumption that function words are frequently grouped together with preceding strong syllables (as in example 4) and not necessarily with the following syntactic head (as suggested in Figure 1).

However, while the parallel approach provides a suggestion for the creation of the lower prosodic domains (foot and prosodic word structure), no suggestion is made for the formation of the higher domains (phonological phrase and intonational phrase); nor do Lahiri and Plank explicitly exclude the influence from other modules of grammar. Furthermore, it has long been part of the indirect reference tradition that it is “crucially [...] not the case that all syntactic boundaries of a certain type must correspond to prosodic boundaries of a given type and vice versa” (Frota 2012: 256). Most researchers assume “prosodic restructuring” (Nespor & Vogel 1986: 172) based on, for example, the type of word (function word vs. lexical word), the size of the phonological phrase, or the amount of recursive nesting.⁷ The indirect reference approaches thus do not take prosodic constituency to be a simple derivative of syntax, but assume that prosodic structure is also formed by means of syntax-independent constraints, among them the rhythmic constraints proposed by the parallel approach.

The difference between these two approaches to prosody and its interfaces necessarily reflects two distinct views of grammar in general. As both directions have been pursued in the proposals made in LFG, the following section briefly discusses how prosodic structure is integrated into the overall grammar architecture and how the indirect reference approach and the parallel approach differ with respect to the communication at the interfaces.

3 Prosody in LFG’s grammar architecture

Several proposals have been made with respect to LFG’s grammar architecture (see also Belyaev forthcoming(b) [this volume]) and a closer discussion of the different approaches to prosody and its interfaces presented in this chapter provides interesting insights into the positioning of the different modules on the one hand, and a general understanding of grammar on the other hand.

Prosodic structure is especially interesting as it is usually taken to represent FORM and is thus placed at one ‘end’ of the FORM-MEANING relationship as, for

⁷For prosodic restructuring mechanisms/‘prosodic markedness constraints’ as expressed in Optimality Theory, see Selkirk (2011: 468ff) for an overview.

example, discussed in Kaplan (1987: 362). In Kaplan’s original proposal, shown in Figure 2, p-structure was not yet part of the grammar architecture; instead, the (word) string was taken to be the external form of the sentence. This is also reflected in Asudeh (2009: 110), where the string is understood as a ‘representation of linear phonology’ and as the center of the syntax–phonology interface, a proposal that has concretely been pursued in the approaches of Bögel et al. (2009, 2010), Dalrymple & Mycock (2011), and Mycock & Lowe (2013) (see Section 4).

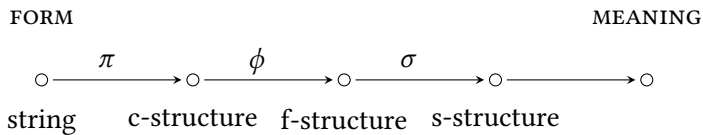


Figure 2: (Simplified) FORM-MEANING relationship in Kaplan (1987: 362)

Proposals in the wider literature have assumed a slightly more complex representation of FORM. Very early, Selkirk (1984) proposed that syntactic structure is first mapped to a phonological (including prosodic) representation, which is then further processed by means of phonological rules and constraints before being mapped to a phonetic representation. In this model the string is not placed between the syntactic and the phonological module, but is the output of the phonological and the phonetic modules. Such a model is very much in line with psycholinguistic models of speech production and comprehension, e.g., as found in Levelt (1999) and in Jackendoff (2002)’s work on Parallel Grammar; see Varaschin forthcoming [this volume].

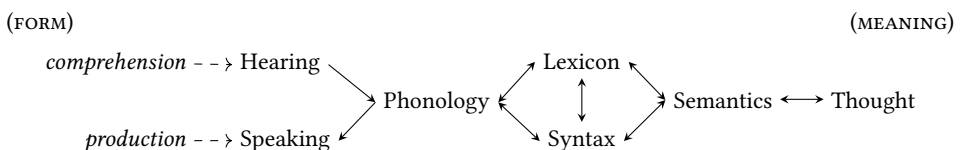


Figure 3: The language processor (cf. Jackendoff 2002: 197, modified)

The model in Figure 3 states clearly what is only implicitly expressed in theoretical LFG:⁸ The different modules, placed between FORM and MEANING, assume a certain directionality, generally termed as ‘comprehension’ (parsing) and ‘production’ (generation) in the wider literature. This is seemingly in conflict with

⁸For example, in Figure 2 by means of arrows, and more explicitly in the pipeline architectures of the numerous computational LFG grammar implementations (see Forst & King forthcoming [this volume]).

the assumption that the different modules exist in ‘parallel’ in LFG (Dalrymple et al. 2019: 265); however, as Jackendoff explicitly remarks, this is not necessarily a hindrance:

P[arallel] A[rchitecture] is nondirectional, but its constraints can be implemented in any order suited to particular processing tasks. (Jackendoff 2010: 589)

‘Parallel’ under this approach refers to the general understanding that each module is subject to its own principles and constraints (= modularity). It does not mean that each component builds a completely isolated structure which then has to be aligned to the output of other modules. Instead, the individual constraints should be adjusted to the processing task at hand (which is either from FORM to MEANING (comprehension/parsing) or from MEANING to FORM (production/generation)).

While this distinction might not carry much weight if a linguistic analysis is provided within one module of grammar (e.g., a syntactic phenomenon), it is crucial when modelling constraints at an interface, as the involvement of two (or more) modules always involves a ‘direction’. The assumptions made by Selkirk above, for example, are made from the perspective of production, while the architecture proposed by Kaplan in Figure 2 (and in general the vast majority of LFG-related linguistic analyses) is made from the comprehension perspective. The acknowledgement of this bidirectionality as made explicit in Figure 3 is fundamental for the discussion of any interfaces between different modules, and thus essential for the proposals on the integration of prosody and its interfaces into LFG.

Models which follow the parallel approach to prosody and its interfaces as detailed in Section 2.2 by assuming that modules are built up independently of each other and that their output is matched for the best alignment at each interface might seemingly be in line with the concept of modules existing in parallel. However, such models are not built to reflect the processing of a given speech signal to understand its meaning (→ comprehension), or the production of a signal expressing a specific thought (→ production).

4 LFG approaches to prosody and its interfaces

With respect to prosody and its interfaces, both the indirect reference and the parallel approach have been explored within the LFG community, mostly with a directional perspective. As these proposals frequently influence each other and

furthermore represent very different views of grammar, the following section provides a chronological overview of the different approaches with a specific focus on the architectural assumptions behind each proposal.

4.1 From c-structure to p-structure: Butt & King (1998)

Butt & King (1998) were the first to introduce a discussion of the syntax-prosody interface and p-structure to LFG. They assumed a mutually constraining model where d(iscourse)- and p-structure are projected off c-structure (in parallel to f-structure), as shown in Figure 4.

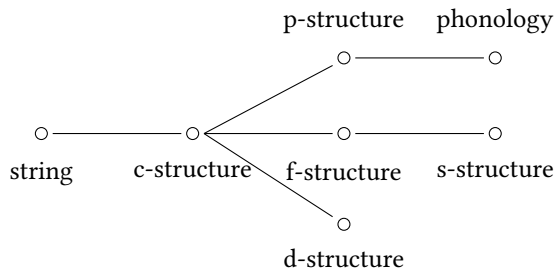


Figure 4: Grammar architecture according to Butt & King (1998: modified)

Under this approach, c-structure is a pivot point between d- and p-structure.⁹ P(rosodic)-structure is viewed as an intermediate between c-structure and the phonological component itself which also contains postlexical phonological rules.

Based on work by Hayes & Lahiri (1991), Butt and King focus on syntactically ambiguous sentences in Bengali, such as example (5).

- (5) ami bHut dekh-l-am
 I ghost see-PST-1SG
 a. ‘I was startled’ (idiomatic)
 b. ‘I saw a ghost.’ (transitive)

Following findings discussed in Hayes and Lahiri, Butt and King assume that prosody can be applied to differentiate between the idiomatic and the transitive interpretation. For Bengali, the assumption for the syntactic-prosodic constituent mapping is that every clause is mapped to an Intonational Phrase (ι , IntP),

⁹In contrast to this chapter which takes prosody to be part of phonology (see footnote 1), Butt and King differentiate between a p(rosodic)-structure and a phonological component.

every NP to a Phonological Phrase (φ , PhP), and every main V or complex predicate is phrased separately. Figure 5 shows the c-structures for (5) and the phrasing possibilities for *bHut dekHlam* which consist of separate phonological phrases in the transitive, and a single phonological phrase in the idiomatic reading.

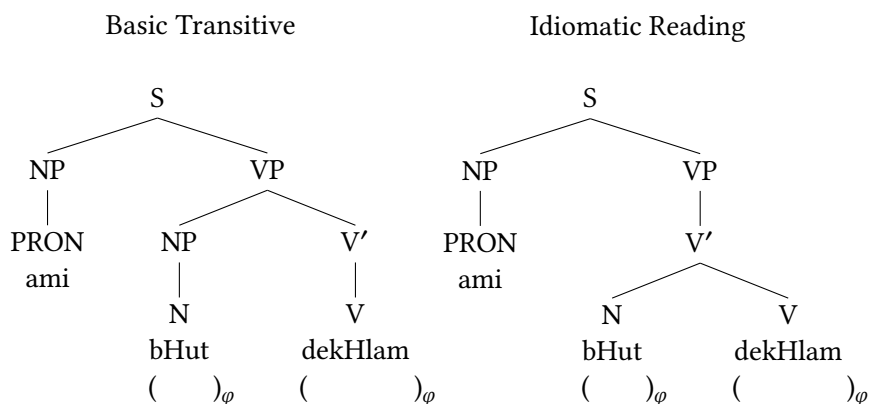


Figure 5: Two c-structure analyses for example (5), Butt and King (1998, modified)

Butt & King (1998) represent p-structure as an AVM structure based on the prosodic hierarchy as shown in (1) above. The AVM structure allows for the inclusion of more detailed information beyond the prosodic domain and the p(honological)-form, such as pitch accents or boundary tones. Butt and King also discuss the linearity issue given with any AVM approach. In order to apply phonological and phonetic processes, it is necessary to preserve the linear order of the string. For a possible solution to this issue, Butt and King point towards projection precedence (Zaenen & Kaplan 1995), which arranges the attributes in p-structure similarly to the string.

Figure 6 shows the prosodic structure of the idiomatic reading in (5), where *bHut* and *dekHlam* are phrased into one phonological phrase (DOM(AIN): P-PHRASE). The AVM includes all the information in the tree structure and the additional information known about the tones in the language, for example, that in a neutral (non-phonological) focus construction, a high tone is associated with the left p-word in the rightmost p-phrase and the whole clause receives a low boundary tone. This information is stored in the AVM (TONE HIGH), but the final association of the high tone with the correct p-form (and the correct syllable in this p-form) is left to the phonological component itself. The reason is that the final p-phrase can only be identified once prosodic phrasing is complete and that the placement

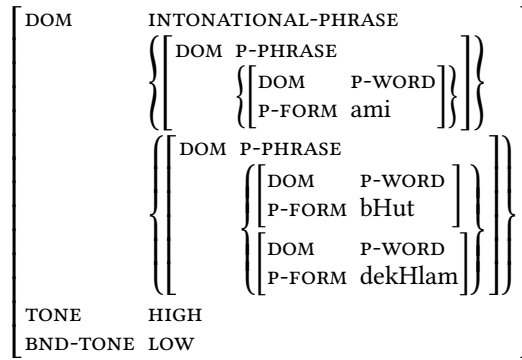


Figure 6: Prosodic structure relating to the idiomatic reading in example (5), neutral focus

of the pitch accent on the correct syllable is solely depending on the phonological structure of the word itself.

Contrastive focus, on the other hand, is indicated by a low pitch accent and a high (intermediate) boundary tone at the level of the phonological phrase, as shown in Figure 7. As the target of the contrastive focus is determined by grammar (here d(iscourse)-structure), the associated pitch accent and boundary tone can be mapped to p-structure together with their domain.

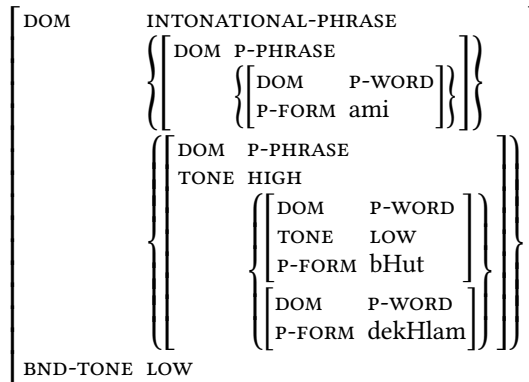


Figure 7: Prosodic structure relating to the idiomatic reading in example (5), contrastive focus

Butt and King determine the tone distribution by using c-structure as a pivot between d-structure and p-structure, as shown in (6).¹⁰

- (6) $(\downarrow_d \text{ FOCUS-TYPE}) =_c \text{ CONTRASTIVE}$
 $(\uparrow_p \text{ TONE}) = \text{ HIGH} \rightarrow \text{ phrasal high}$
 $(\downarrow_p \text{ TONE}) = \text{ LOW} \rightarrow \text{ local low}$

The approach proposed by Butt & King (1998) was later taken up in Bögel et al. (2008) in their analysis of Urdu *ezafe*. In the *ezafe* construction in (7), the *ezafe* clitic is syntactically grouped with the following modifying noun, but is prosodically attached to the previous head noun.

- (7) sher=e panjAb
lion=Ez Punjab
‘a/the lion of Punjab’
Syntactic Phrasing: [[sher] [e panjAb]]
Prosodic Phrasing: ((sher e) panjAb)

Example (7) shows a typical mismatch between syntactic and prosodic structure, which is difficult to account for if prosodic constituency is directly based on syntactic constituency. The solution proposed in Bögel et al. (2008) integrates the *ezafe* clitic (CL-FORM) into the phonological phrase using a number of bookkeeping features to make sure an *ezafe* clitic is present, as shown in Figure 8.

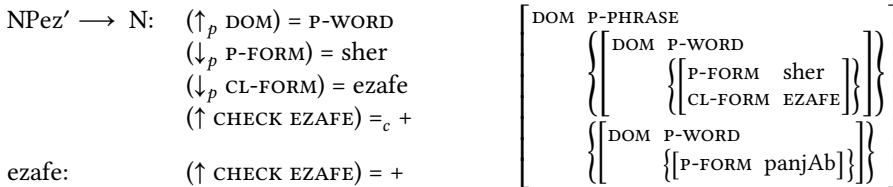


Figure 8: (Reduced) *ezafe* rule and the resulting p-structure in Bögel et al. (2008)

This approach is not entirely satisfactory. For one, in this approach it is actually the noun which is ‘checking’ for a following clitic, instead of the clitic ‘asking’ to be grouped with a preceding prosodic host. Furthermore, this approach does not allow for a language-specific expression of prosodic principles, e.g.,

¹⁰For the interested reader: Focus in Bengali can also be signalled by the clitic -o. Following Lahiri & Fitzpatrick-Cole (1999), Butt and King assume a lexical high tone which is introduced onto the prosodic word with the clitic’s lexical specifications.

the general integration of *enclitics* into the preceding prosodic domain, and of *proclitics* into the following prosodic domain. Instead, individual specifications have to be created for each clitic. This is not only unintuitive, but also does not allow for any predictions to be made about prosodic structure in general.

Summing up, Butt and King make a first proposal to include prosodic information into LFG and show how this can interact with d- and c-structure. In contrast to Hayes & Lahiri (1991)'s original approach (and in contrast to the claim made in Dalrymple et al. 2019), their model does not permit the direct reference of phonological restructuring rules to relations internal to syntactic structure (e.g., to 'right sister' or modifier-head-constructions), but provides an indirect, modular approach to the interface.

Butt and King distinguish between two structures: p-structure and the phonological component. P-structure only includes the information that is pre-determined by other modules of grammar, e.g., pitch patterns introduced by different sentence types and focus, and prosodic constituency based on syntactic constituency. This information serves as input to the phonological component (not further defined in their paper) and its inherent rules and constraints, which include prosodic restructuring, or the placement of pitch accents on the correct syllables within the right domains. This directional analysis from c-structure to p-structure to phonology reflects part of the production process in Figure 3. However, Butt and King's model (as shown in Figure 4) is generally not in line with the architectural assumptions made in Figure 2 and Figure 3 in that the string is not the representative of FORM: neither is the string equal to the phonological output nor is it closely associated with the phonological module. Without this connection, it is unclear how the string could be realised in terms of a (physical) speech signal.

4.2 Prosody and i-structure: O'Connor (2005)

In his thesis, O'Connor (2005) discusses the interface between prosody and information structure. O'Connor (2005) assumes a bidirectional approach, which distinguishes between a 'hearer-based' and a 'speaker-based' approach. He explicitly focusses only on the hearer-based direction (from p- to d-structure → comprehension) and leaves the speaker-based direction (from d- to p-structure → production) to further research.

O'Connor's approach is based on the AM/ToBI framework (see Section 2.1), but the description of accents is restricted to High and Low tones only.¹¹ He is

¹¹O'Connor does not distinguish between different types of pitch accents and how these may relate to specific i-structure categories, e.g., the distinction between broad and narrow focus based on different pitch patterns (a.o., Baumann et al. 2007).

mostly concerned with utterances where a difference in meaning is expressed solely by means of prosodic emphasis (expressed by capital letters in example (8)).

- (8) a. He rode [a green DRAGON]_{foc}.
 b. He rode a [GREEN]_{foc} dragon.

The two propositions have a different information structure. Example (8a) can be the answer for a question with a broad focus (e.g., *Did he ride a green dragon or a thestral?*), while example (8b) is more likely to be the (contrastive) answer to a question like *Did he ride a green dragon or a blue dragon?*

In his proposal, O'Connor assigns a central role to i-structure. As the AM/ToBI system is not concerned with the influence of syntactic structure on prosody, O'Connor assumes that prosody and i-structure can be related to each other without syntactic mediation, as shown in Figure 9.¹²

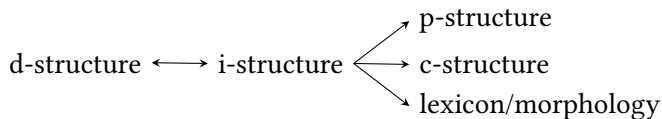


Figure 9: Architecture proposed in O'Connor (2005: Fig 6.3, p. 142, modified)

Following the general idea behind autosegmental approaches (Goldsmith 1976), O'Connor pursues the idea of a representation of tonal information independent from the segmental/phonemic representation. He proposes that p-structure should be represented by a hierarchical constituent structure (thus paralleling c-structure). Via so-called 'tune structure rules' like the ones in (9), a tree-like structure is created to represent intonation where the terminal nodes correspond to underspecified tonal events: t^* represents a pitch accent, t^- a phrase accent, and $t\%$ a boundary tone.

- (9) $n \geq 1$
 a. $TUNE_{IP} \rightarrow \text{tune}_{ip}^n \ t\%$
 b. $\text{tune}_{ip} \rightarrow t^{*n} \ t^-$

¹²O'Connor does not completely exclude the influence of c-structure on prosody, but only acknowledges a relevance of the linear and hierarchical syntactic structure of the clause for the length of the transition between tonal events and the alignment of the pitch in general.

As a result, each prosodic tree constructed on the basis of these rules has four obligatory nodes: the prosodic ‘intermediate phrase’ (tune_{ip}) consists of a nuclear accent t^* and a phrase accent t^- while the prosodic ‘intonational phrase’ (TUNE_{IP}) consists of at least one intermediate phrase and a boundary accent $t\%$. For example (5) (see also Figure 5, basic transitive) from Butt & King (1998), O’Connor proposes the p-structure in Figure 10.

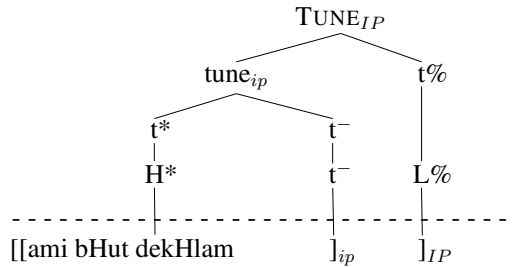


Figure 10: O’Connor’s p-structure as applied to example (5) from Butt & King (1998)

The tree representation is mainly concerned with the organisation of tonal events; the material below the dashed line includes the orthographic tier and the prosodic domain information in the form of bracketing.¹³

O’Connor emphasises the point that under his approach, the association of the High tone is not left to a further phonological component as proposed in Butt & King (1998) and discussed above in Section 4.1. It is however not quite clear how the High tone is associated with the correct string sequence in O’Connor’s approach, as no formal alignment of string and pitch (i.e., c-structure and p-structure) is established in his thesis. Indeed, in the data provided by Butt and King, the High tone should be assigned to the ‘leftmost’ prosodic word in the ‘rightmost’ phonological phrase. As O’Connor collapses all three phonological phrases proposed by Butt and King under one tune_{ip} , it is not clear how the association of the High tone with the correct word can be ensured.

With respect to i-structure, O’Connor assumes that categories like FOCUS and TOPIC are organised linearly in an utterance, and assigns each to one tune_{ip} , as shown in (10). If there are not enough tunes, then the assumption is that there is no topic correspondent.

¹³As O’Connor’s (2005) main focus is on the relation between intonation and discourse functions, the encoding of further prosodic/phonological information, e.g., syllable structure, or lexical stress, is not further discussed in his thesis.

$$(10) \quad \text{TUNE}_{IP} \longrightarrow \begin{array}{ccc} \text{tune}_{ip} & & \text{tune}_{ip} & \text{t}\% \\ \downarrow \in \{\uparrow_d \text{ TOPIC}\} & & \downarrow \in \{\uparrow_d \text{ FOCUS}\} & \end{array}$$

As O'Connor notes, there are a number of cases where the proposed association of i-structure categories and tunes does not work. Sentences like *'It broke.'* will only have one tune, indicating a FOCUS. However, in i-structural terms, *It* is a TOPIC. This mismatch between tunes and i-structure roles is discussed (O'Connor 2005: 161), but not resolved.

In conclusion, O'Connor's indirect, directional approach to the relationship between prosody and i-structure suggests an alternative to the syntactocentric view proposed in Butt & King (1998). However, there are a number of outstanding questions. Besides the incomplete association of i-structure categories to tune-structure rules just discussed, there is also the fundamentally important question as to how tonal events can be associated with their targets without reference to the morphosyntactic string, or how common phenomena, e.g., the (syntactic) scope of a prosodically expressed focus, can be determined without reference to syntactic constituency. The missing association of p-structure, string, and c-structure, and other unresolved questions thus only allow for an analysis of a more descriptive nature.

4.3 The string as an interface between c- and p-structure: Bögel et al. (2009, 2010)

Based on the realisation that the frequent misalignment of prosodic and syntactic constituents would seriously complicate previously established prosody-syntax mapping algorithms, Bögel et al. (2009) pick up on the notion of the parallel approach discussed in Section 2.2. The underlying assumption is that the prosodic component operates independently of syntax and that the two components are not related via LFG's projection architecture. To account for the cases where syntax is influenced by prosody, Bögel et al. (2009) assume a directional 'pipeline' architecture (from the comprehension perspective): First, an independent prosodic component interprets various phonological properties thus establishing the boundaries of prosodic units. This information on prosodic constituency is then made available to syntax by inserting prosodic bracketing features into the terminal string of c-structure.¹⁴

Bögel et al. (2009) discuss a number of different phenomena, among them Urdu *ezafe* (Bögel et al. 2008: see Section 4.1). They extend the c-structure rules by

¹⁴Under this approach, the string has a central role as it includes information from both the syntactic and the prosodic component (similar to the understanding of the string in Asudeh 2009: 110 as a 'linear representation of phonology').

adding left and right prosodic brackets ('lexical categories' RB and LB) which reflect prosodic constituency.

- (11) a. EzP → EZ RB N
 b. NPez' → LB [...] N

The inclusion of brackets greatly simplifies the rule originally used in Bögel et al. (2008: Figure 10) where a number of CHECK-features were applied to control for an *ezafe* clitic following the head noun. The resulting c-structure representation, shown in Figure 11, allows for the depiction of the misalignment between syntactic and prosodic structure.

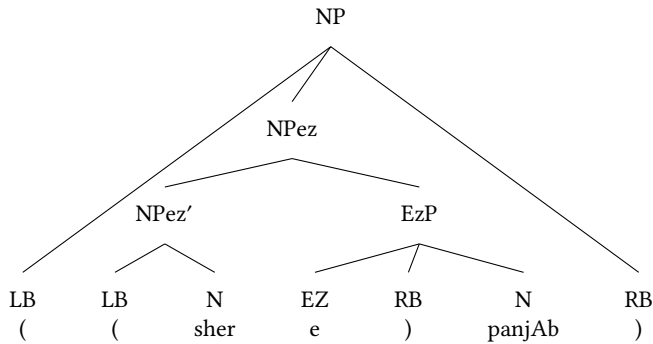


Figure 11: Urdu *ezafe* analysis as proposed in Bögel et al. (2009)

Another aspect discussed in Bögel et al. (2009) is the prosodic resolution of syntactically ambiguous structures. Consider the following example, where *old* can either modify only the first noun ((12a)) or scope over the whole coordination ((12b)). Each possibility is accompanied by a distinct prosodic grouping.

- (12) a. [old men] and [women]
 (old men) and (women)
 b. [old [men and women]]
 (old (men and women))

The paper postulates a 'Principle of Prosodic Preference', according to which the syntactic component disprefers syntactic structures whose constituent boundaries do not coincide with prosodic boundaries. For the implementation, Bögel et al. (2009) use a metarule, which systematically transforms the rules of the syntactic component. In the following metarule, CAT is a nonterminal category,

and RHS denotes the regular language over categories which are annotated with co-describing constraints.

(13) $CAT \rightarrow RHS$

In Bögel et al.'s metarule in (14), the top part of the rule will match a (recursive) sequence of CAT surrounded by prosodic brackets (LB and RB). The bottom part will match the RHS regular expression if all occurrences of LB or RB are ignored, thus preventing the inserted prosodic brackets from ruling out a valid syntactic analysis.¹⁵

(14) $CAT \rightarrow LB\ CAT\ RB$
 $| RHS / [LB | RB]$
 Disprefer

The Principle of Prosodic Preference is enforced via the 'Disprefer' optimality mark,¹⁶ which assigns a dispreference mark to the construction every time the bottom part of the metarule in (14) is applied. In the case of several possible syntactic analyses (as in (12)), this extension effectively ranks the analyses: The top half of the rule only applies if the prosodic brackets match the syntactic structure, while the syntactic analyses with no matching prosodic brackets will be parsed by the bottom half of the rule, but will receive a 'Disprefer' mark. This allows for constructions with matching prosodic and syntactic brackets to be preferred, while constructions with non-matching brackets will only be valid if a preferred solution (with matching brackets) is not available.

This first approach to the interface was extended in Bögel et al. (2010) which discusses second position (2P) clitics in Russian and Serbian/Croatian/Bosnian (SCB). It is concerned with examples like (15), where a clitic cluster (CCL) disrupts the NP *Taj čovek*.

(15) [Taj *joj ga je* čovek] poklonio.
 that her it AUX man presented
 'That man presented her with it.' (Schütze 1994)

These clitics appear in the second position after a first prosodic word without regard to syntactic requirements ((16)).

¹⁵The 'Ignore operator' / was first introduced in Kaplan & Kay (1994).

¹⁶See Forst & King forthcoming: section ?? [this volume] for a description of optimality marks and relevant references.

- (16) $((\text{Taj})_\omega (\text{joj ga je})_{cl})_\omega (\text{čověk})_\omega)_\varphi (\text{poklonio})_\varphi$
 That her it AUX man presented
 ‘That man presented her with it.’

Such a structure is problematic for traditional LFG accounts, because it is difficult to account for the clitics’ appearance within an NP and to furthermore retrieve the clitics’ functional contribution to the clausal f-structure. With a purely syntactic account, this information is locked into the NP’s f-structure.

Bögel et al. (2010) resolve this issue by assuming a shared responsibility between the syntactic and the prosodic component: While the syntactic component ensures the availability of the functional information by placing the clitic in the (linear) first position, the prosodic component ensures the correct position of the clitics within the clause and places the clitics following the first prosodic word. This prosodic repair mechanism has been shown to apply crosslinguistically and was dubbed ‘prosodic inversion’ by Halpern (1995).¹⁷

In order for the clitic to appear in the correct syntactic position, Bögel et al. (2010) define the rule in (17), where RHS_S denotes the possible expansion of the clausal S node with left and right brackets (as discussed above). LB_S is a pre-terminal node that marks the left edge of a clause and allows syntactic/prosodic constraints to be aligned with respect to clause boundaries. CCL can optionally appear as a prefix to the S expansion; the $\uparrow=\downarrow$ annotation ensures the processing of the clitics’ clause-level functional information.

- (17) $S \rightarrow \text{LB}_S (\text{CCL}) \text{RHS}_S$
 $\uparrow=\downarrow$

To account for the prosodic placement, Bögel et al. (2010) distinguish between a prosodic and a syntactic (c-structure terminal) string which includes the lexical formatives discussed above. The interface between these (usually aligned) strings is a regular relation, where the syntactic string is the ‘upper language’ and the prosodic string is represented by the ‘lower language’. In the simplified illustration in (18), the upper language/syntactic string clitic sequence (CS:0) immediately following the clause boundary (ς is placed after the first prosodic word ω in the prosodic string in the lower language/prosodic string (0:CS).

- (18) $\frac{\text{s(yntactic)-string ('upper')}: (\varsigma \text{ CS } \omega \ 0)}{\text{p(rosodic)-string ('lower')}: (\varsigma \ 0 \ \omega \ \text{CS})}$

¹⁷For further work in LFG, see an account of prosodically determined second position clitics in Vafsi in Bögel et al. (2018).

The regular relation has the effect that strings with syntactically clause-initial clitic sequences are related to strings where those clusters appear on the other side of an adjacent prosodic word. The sentence-initial position allows for the functional information to be made available to syntax, but violates the prosodically dependent clitic's need for a preceding host. The second position in the prosodic string satisfies this prosodic constraint in that the clitic is placed following a valid prosodic host.

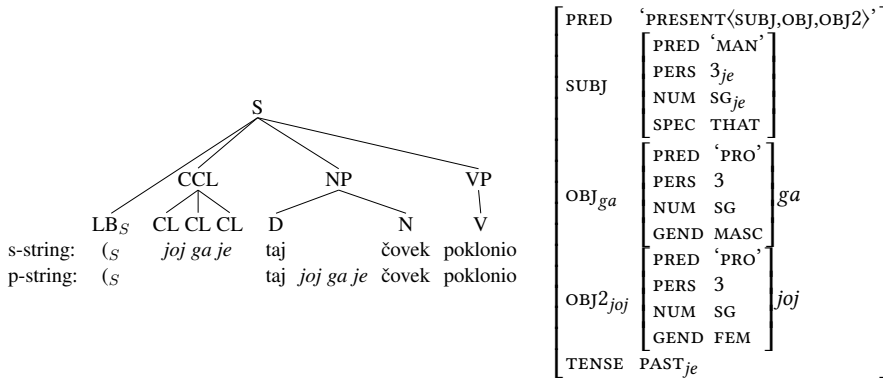


Figure 12: 2P clitics and the syntax-prosody interface (Bögel et al. 2010: 13)

In conclusion, Bögel et al. (2009) set up a directional architecture where they assume an independent prosodic component to process information related to prosody. Information on prosodic constituency is only entered into the terminal syntactic string in the form of lexical formatives, that is, the approach pursues a softer version of the parallel approach discussed above in Section 2.2. This allows for an account of general prosody-syntax misalignment, but also for a prosodic resolution of syntactically ambiguous structures. The approach was furthermore extended to account for second position clitics as well.

There are also several problems with this approach. For one, no further prosodic information can be transmitted via the string. However, information on pitch patterns, accent types or other prosodic features can be crucial for a correct interpretation of meaning.

Another factor is the relatively coarse-grained representation of prosody. Prosodic units can be relevant down to the syllable, a granularity that becomes difficult to represent by interspersing brackets with the terminal c-structure string. And finally, integrating the prosodic structures into the c-structure analysis makes the representation not only difficult to process, but is also questionable in terms of modularity.

4.4 A strictly parallel approach: Dalrymple & Mycock (2011), Mycock & Lowe (2013)

Dalrymple, Mycock, and Lowe base their approaches¹⁸ on the assumption that the prosodic and the syntactic component are parallel but separate components, which goes beyond the distinction between direct and indirect approaches to prosody and its interfaces as briefly discussed in Section 2.2. In the parallel approach, syntactic structure has no influence on the formation of prosodic structure (and vice versa).¹⁹ Instead, each structure is built up independently: syntactic structure as traditionally assumed, and prosodic structure based on rhythmic principles, more specifically, the trochaic foot. In contrast, the indirect reference approach assumes that rhythmic structure is only one factor among many which contribute to the formation of prosodic structure.

The approach represents prosodic constituency in a tree-like structure, assuming the constituents proposed by Selkirk (1995) (see Section 2.1). Similar to the proposal made in Bögel et al. (2010), the interface between the syntactic and prosodic components is the interface between a s(yntactic)-string and a p(honological)-string. The “linguistic signal” (Dalrymple et al. (2019: 407), the nature of which is not further defined) is parsed into minimal syntactic units in the s-string, and into minimal prosodic units (i.e., syllables) in the p-string. A (very simplified) representation of the example sentence *Anna was studying at the university* is shown in Figure 13 (see Figure 15 for a complete picture).

Figure 13 displays the syntactic component (s-string and c-structure) in the top part. The bottom part represents p-structure: the p-string and the prosodic tree. The p-string is parsed into syllables (but see below for further specifications) which are grouped into prosodic words. Following Lahiri & Plank (2010), prosodic structure is built based on rhythmic principles, specifically on the trochaic foot (see Section 2.2). The representation omits the foot structure, but the underlying formation algorithm is still visible in the fact that the left edge of each prosodic word is placed with the syllable carrying primary stress in a lexical, (syntactic) word; e.g., u.ni.(VER.si.ty)_ω. Function words and feet built on secondary stress (e.g., (,u.ni)_{f_t}) seem to generally be phrased with the preceding

¹⁸The approach to the interface described in this section was developed in a number of works, namely Mycock (2006); Dalrymple & Mycock (2011); Mycock & Lowe (2013) (see also Lowe 2016 and Jones 2016 for further discussion); the version described here is part of the prosody chapter in Dalrymple et al. 2019.

¹⁹Dalrymple et al. (2019: 398) classify their approach as indirect reference. However, all indirect reference approaches include syntactic structure as a main factor for building up prosodic structure (mostly from the perspective of production). This is not the case here.

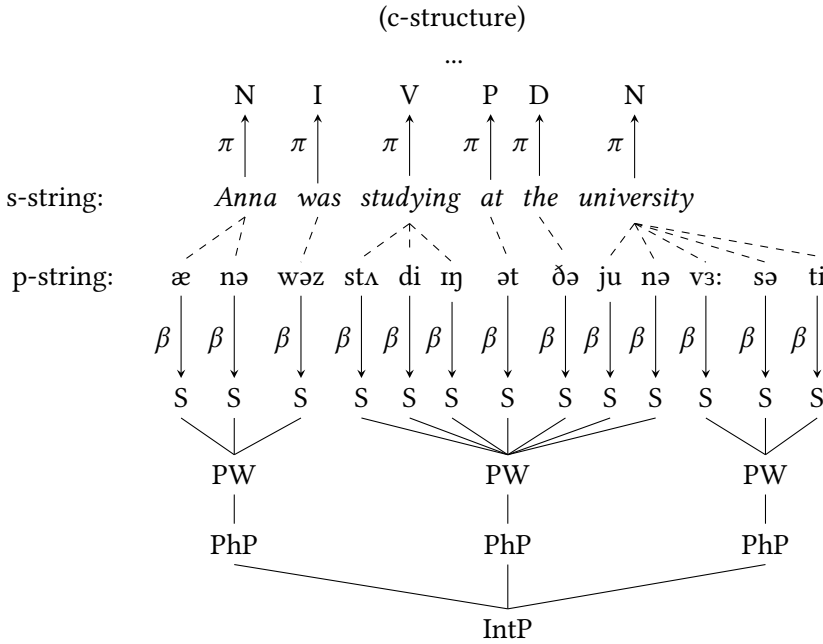


Figure 13: Simplified interface for *Anna was studying at the university* (Dalrymple et al. 2019: 408, showing only c-structure terminal nodes)

prosodic word.²⁰ The following example shows the prosodic phrasing according to a trochaic foot structure, with primarily stressed syllables in capital letters.

- (19) (Anna was) (STUDying at the uni) (VERsity)

The formation of prosodic words based on rhythmic principles naturally leads to regular mismatches between syntactic and prosodic units. However, the approach raises the question whether these units are indeed prosodic words or whether they should rather be defined as phonological phrases. If these units are prosodic words, then the question arises how phonological phrases are defined under this approach. In Figure 13, each phonological phrase is identical with a prosodic word, which is a crosslinguistically very unusual 1-1 relationship.²¹

²⁰It would be interesting to see how this approach can be applied to cases where the first syllable of a prosodic unit is unstressed as in a modified version of example (19), *Anna or Ravi and Karla* ... (Anna ∨ (Ravi ∧ Karla)), where the prosodic boundary is realised directly after *Anna* (Wagner 2010), while the rhythmic approach would predict the prosodic boundary to occur after *or*.

²¹See, for example, the family of BINMIN-constraints, which require for a higher prosodic domain to contain more than one unit of a lower prosodic domain (a.o., Ghini 1993; Inkelas & Zec 1995).

Mycock & Lowe (2013) extend the string interface by assuming that the string-units are not atomic but should rather be seen as feature bundles, represented as AVMs. The relation between the two strings and their units is regulated through information stored in the lexicon. While the lexical s(yntactic)-form contains the traditional morphosyntactic information, the p(honological)-form contains information on segments and syllable structure as well as the feature SYLLSTRESS which indicates the primary lexical stress position.

In addition to the lexical information, the feature structures at the interface also include information on the edges of constituents in the respective modules. These ‘edge features’ are necessary to allow for the matching of prosodic and syntactic constituents, e.g., in order to prosodically disambiguate syntactically ambiguous structures. Mycock and Lowe define a number of mechanisms to make the edges available to the strings: \swarrow and \searrow for the left and right edges of syntactic nodes, and \swarrow and \searrow for the left and right edges of prosodic nodes.²² Figure 14 shows the AVMs for the first syntactic and the first prosodic unit of example (19), where the values of l(ef) and r(ight) consist of a set of syntactic and prosodic nodes whose edges are represented by this particular form.

s-string unit <i>Anna</i>	p-string unit <i>æ</i>
$\left[\begin{array}{l} \text{FM ANNA} \\ \text{L } \{\text{IP, NP, N}\} \\ \text{R } \{\text{NP, N}\} \end{array} \right]$	$\left[\begin{array}{l} \text{FM } \text{æ} \\ \text{SYLLSTRESS P} \\ \text{L } \{\text{INTP, PHP, PW}\} \\ \text{R } \{\} \end{array} \right]$

Figure 14: Feature structure for first unit of the p-string and the s-string (Dalrymple et al. 2019: 412)

At the interface, a “Principle of Interface Harmony” ensures that the best-matching parses between the p-string and the s-string are preferred. Note, however, that the approach does not explain how this preference is implemented.²³

Furthermore, the question of which syntactic and prosodic constituent edges should be matched, that is, which prosodic boundary type is important to syntax and vice versa, is left for future research (Dalrymple et al. 2019: 419). This is surprising, given the extensive existing literature on the topic, but in a sense it is a necessary consequence of assuming strictly parallel modules as has been discussed in Section 2.2.

²²See Dalrymple et al. (2019) for the exact definitions.

²³Lowe (2016) presents a possible implementation of the Principle of Interface Harmony using additional formal power in form of OT constraints (see also Lowe & Belyaev 2015). A critical discussion of this approach can be found in Bögel (2015: Ch.6).

Apart from the interface to syntax, the string interface also serves as an interface between prosody and semantics and i-structure.²⁴ The semantics-prosody interface is demonstrated by means of declaratives, where the intonational contour distinguishes between declarative statements and questions. In order to make the semantic information available at the string interface, the c-structure receives a ‘label’ *PolarIntSem* along with the meaning constructor [*PolarInt*]. Similar to the edge values, this label is handed down to the s-string where it is placed in the rightmost AVM.

For the prosodic interpretation of a declarative question, information on pitch is required. This information is included in the form of H and L pitch accents and boundary tones.

Dalrymple et al. (2019) assume that in English declarative questions, a nuclear L tone is associated with the stressed syllable of the first prosodic word in the last phonological phrase, and an H boundary tone at the right edge of the Intonational Phrase.²⁵ Similar to O’Connor (2005) above (Section 4.2), they annotate prosodic structure by means of prosodic phrase structure rules. In addition, a label *PolarInt* appears at the rightmost AVM of the p-string. The rule in (20) can then be read as follows: In this phonological phrase, assign a nuclear tone L to the leftmost syllable with primary stress (σ^s) and a right boundary tone to the rightmost unit;²⁶ if these constraints are satisfied, create a label *PolarInt* which appears as a set member of the rightmost unit’s right edge.

$$(20) \text{ IntP} \longrightarrow \text{PhP}^* \quad \text{PhP} \\ ((\sigma^s \text{ N_TONE}) = \text{L}) \\ (\tau_{\downarrow} \text{ RB_TONE}) = \text{H} \Rightarrow \\ \text{PolarInt} \in (\tau_{\downarrow} \text{ R})$$

Figure 15 shows part of the full analysis for example (19), the PP *at the university*. In addition to the edge features, both labels, *PolarIntSem* and *PolarInt* appear at the right edge of the string interface. The Principle of Interface Harmony requires both labels to co-occur for the overall structure to be grammatical, but the matching process is not further detailed here.

²⁴As the description of the interface to i-structure is similar to the one provided for the prosody-semantics interface, the interested reader is referred to Dalrymple et al. (2019) for details.

²⁵The prosodic expression of declarative questions in English shows much more variability than assumed here, see, e.g., Gunlogson (2003) for discussion of different contours.

²⁶The assignment of a right boundary tone here does not distinguish between ‘boundary tones’, e.g. H%, which appear as boundary tones of intonational phrases, and ‘phrase accents’, e.g., H⁻, which appear at the edges of phonological phrases. If both edges fall together, these tones form combinations, e.g. H-H%, which can be crucial for an interpretation. The use of a boundary tone with a phonological phrase unfortunately collapses this distinction.

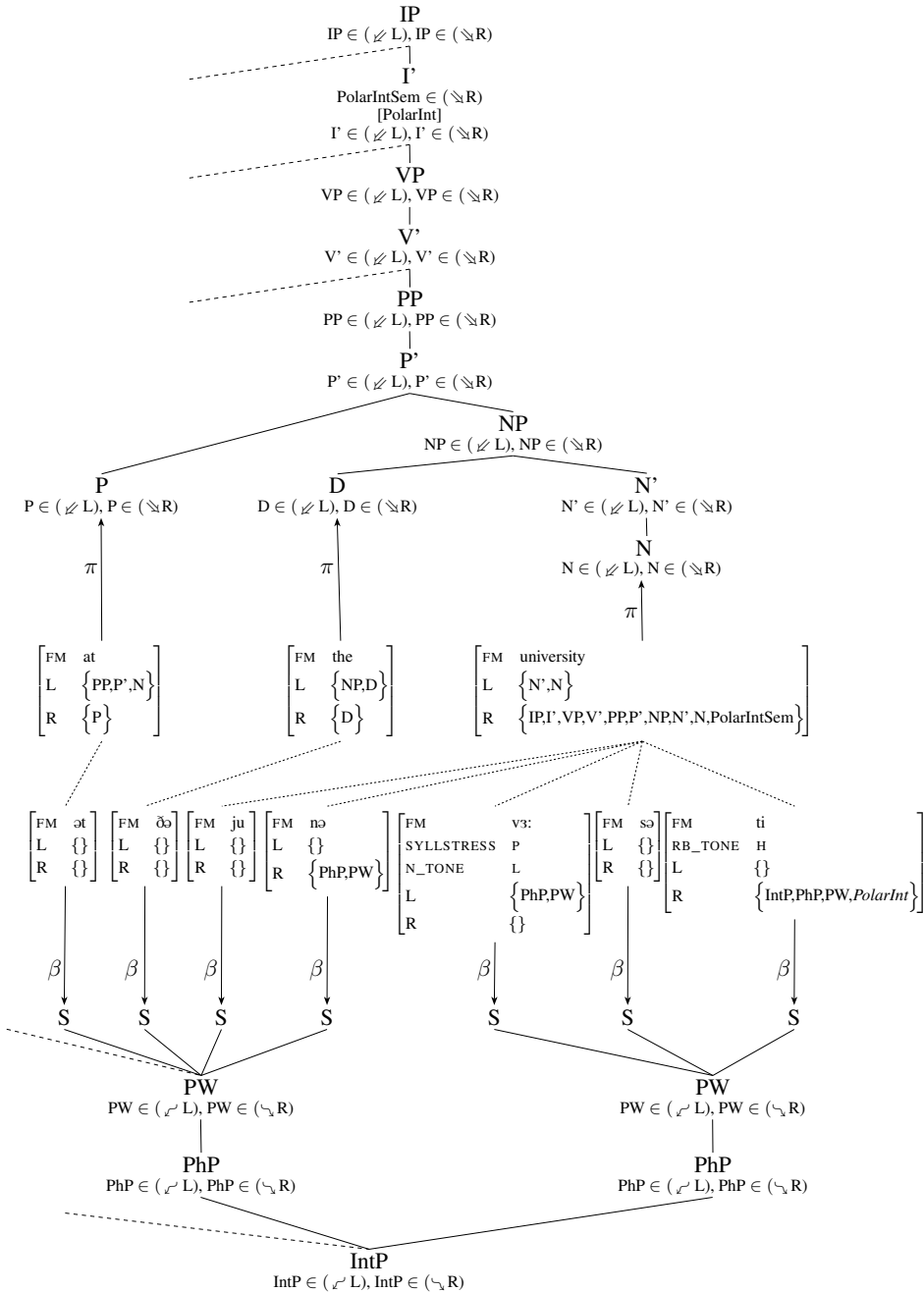


Figure 15: Analysis of the PP in the declarative question *Anna was studying at the university?*

In conclusion, the non-directional model proposed by Dalrymple et al. (2019) pursues the idea of modularity in the extreme: the syntactic and the prosodic component are taken to be completely independent structures which do not allow for any co-description mechanisms as commonly found in LFG (see Belyaev forthcoming(a) [this volume]). Instead, information about (at least) syntactic and prosodic constituency, semantics, i-structure, and intonational contours is handed to the respective strings. The interface between the syntactic and the prosodic component is then situated between the s-string and the p-string, where matching edges and ‘labels’ are preferred according to the Principle of Interface Harmony.

Apart from initial suggestions involving OT-constraints in Lowe (2016) and Lowe & Belyaev (2015), the formal implementation of the Principle of Interface Harmony is not further detailed. Given that there are numerous combinatorial possibilities of prosodic constituents, pitch accents, phrase accents, and boundary tones, and hardly any of them can be mapped to one particular interpretation, but are always co-dependent on other modules of the grammar, the matching of labels at the string interface will most likely prove to be difficult. The introduction of these labels and the mingling of information from different modules is, however, a necessary consequence of the parallel approach. The reduction of the interface to the strings implies that *all* potentially relevant information from other modules has to be duplicated and appear as part of the string where it might or might not be matched against the material in the parallel string. As it was the case in Bögel et al. (2009, 2010) (Section 4.3), this is also problematic with respect to modularity.

The extensive duplication and blending of structures can be avoided by assuming a more traditional co-descriptive approach, while at the same time acknowledging modularity in that each module only processes information related (i.e., ‘native’) to its module. This indirect reference approach was first pursued in Butt & King (1998), and was further developed in Bögel (2015) and subsequent work as discussed in the next section.

4.5 Production and comprehension: Bögel (2015)

Starting with her dissertation in 2015, Bögel developed a directional indirect reference model of the prosody-syntax interface that enables the integration of a speech signal into LFG and can account for a vast variety of phenomena from both perspectives, production and comprehension. In this approach, illustrated in Figure 16, the interface between c-structure and p-structure is regulated via two transfer processes, the ‘transfer of vocabulary’ (ρ), which exchanges phonological and morphosyntactic information of lexical elements via the multidimen-

sional lexicon, and the ‘transfer of structure’ (\mathfrak{H}), which exchanges information on syntactic and prosodic phrasing, and on intonation.

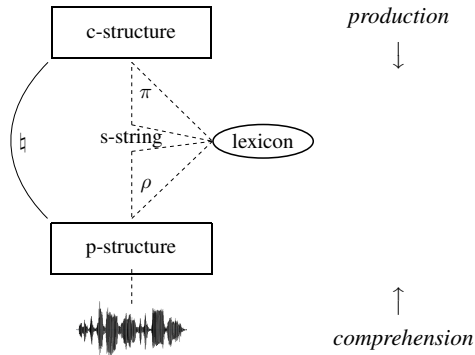


Figure 16: Abstract overview of the prosody-syntax interface (Bögel 2015)

The transfer of vocabulary requires a lexical entry to contain detailed information on (at least) the morphosyntactic as well as the phonological form (s-form and p-form, Dalrymple & Mycock 2011). Following Levelt et al. (1999), Bögel (2015) develops a multidimensional lexicon where the s-form encodes the traditional morphosyntactic information, and the p-form contains information on segments, the metrical frame and prosodic word status.²⁷ Figure 17 shows the (shortened) lexical entries for *čovek* (‘man’), *taj* (‘that’), and *joj* (‘her’) from example (15) (repeated in (21)) from SCB, where the clitic cluster (*joj ga je*) is placed in the prosodic second position, syntactically ‘interrupting’ the NP *taj čovek*.

- (21) [Taj *joj ga je* čovek] poklonio.
 that her it AUX man presented
 ‘That man presented her with it.’ (Schütze 1994)

The lexical p-form entries of *čovek* and *taj* are both marked as full prosodic words (ω). In contrast, the p-form of *joj* is marked as a prosodically deficient enclitic ($=\sigma$), that is, it is prosodically dependent on a preceding host. Following the concept of modularity, each dimension can only be accessed by the related module: c-structure can access the s-form and p-structure the p-form. However, during the transfer of vocabulary, the lexicon also assumes a ‘transducer function’ between s-form and p-form: If a particular dimension is accessed (e.g., s-form from

²⁷ A third dimension, ‘concept’, which includes semantic information is assumed as well, but not discussed further here.

s(yntactic)-form			p(honological)-form	
čovek	N	(↑ PRED) (↑ PERS) ...	= 'čovek' = 3	P-FORM [tʃovek] SEGMENTS /tʃ o v e k/ METR. FRAME (σσ) _ω
taj	PRON	(↑ PRED) (↑ PRON-TYPE) ...	= 'pro' = demon	P-FORM [taj] SEGMENTS /t a j/ METR. FRAME (σ) _ω
joj	PRON	(↑ PRED) (↑ PRON-TYPE) ...	= 'pro' = pers	P-FORM [joj] SEGMENTS /j o j/ METR. FRAME =σ

Figure 17: Lexical entries for SCB *čovek* ‘man’, *taj* ‘that’, and *joj* ‘her’

c-structure), the associated dimensions become available as well and the information stored in them is projected to their respective structures (e.g., p-form information becomes available to p-structure).

P-structure itself is represented by the p-diagram, a compact linear representation of the utterance. The p-diagram is structured syllablewise, where each syllable is part of a vector (v(ECTOR) INDEX) which associates the syllable with relevant segmental and suprasegmental phonological information.²⁸ During the transfer of vocabulary, the information stored with each lexical item’s p-form is stored in the p-diagram. Figure 18 illustrates this process for example (21).

↑ PHRASING	↑ =σ	=σ	=σ	(σ) _ω	(σ	σ) _ω	...	↑
LEX.STRESS	-	-	-	prim	prim	-	...	
SEGMENTS	/joj/	/ga/	/je/	/taj/	/tʃo/	/vek/	...	
V. INDEX	S ₁	S ₂	S ₃	S ₄	S ₅	S ₆	...	

Figure 18: The p-diagram with material from the transfer of vocabulary from example (21) (production)

The p-diagram’s content in Figure 18 is identical with the lexical p-form information in Figure 17: *tʃovek*, for example, consists of two syllables, each of which contains a number of segments. The first syllable has primary stress and the complete word forms a prosodic word. These attributes and their values are stored

²⁸The choice of the underlying prosodic or phonological unit and of the different attributes is up to the researcher.

for each syllable, thus creating a linear representation of the phonological string, with a vertical representation of the different values associated with each part of the phonological string.

In addition to the lexical information, p-structure receives information on syntactic constituency through the transfer of structure. This approach assumes match theory²⁹ (see Section 2.2), where each syntactic clause is mapped to an intonational phrase and each XP is mapped to a phonological phrase. The c-structure annotation in (22) models this approach to the mapping between syntactic and prosodic constituents for the clausal node S.

$$(22) \quad \begin{array}{c} S \\ (*(T(*)) S_{min} \text{ PHRASING}) = {}_l(\\ (*(T(*)) S_{max} \text{ PHRASING}) =)_r \end{array}$$

This annotation can be read as follows: Take all terminal nodes (T) of the current node (*, here S), for the attribute PHRASING assign a left IntP boundary (${}_l()$) to the leftmost syllable (S_{min}) and a right IntP boundary to the rightmost syllable (S_{max}) in p-structure. The transfer of structure thus encodes information on larger prosodic domains in p-structure. Taken together, the transfer of vocabulary and the transfer of structure thus provide an initial input to p-structure based on lexical phonological information on the one hand, and on syntactic constituency in form of larger prosodic domains on the other hand (see Figure 19).

As discussed above in Section 4.3, the syntactic analysis of example (21) positions the clitic cluster in the sentence-initial position: [*joj ga je*]_{CCL} [*taj čovek*]_{NP} [*poklonio*]_{VP}. As information is accumulated, a prosodic constraint violation becomes apparent (which, in line with modularity, syntax neither recognized nor cared about): The clitics are placed in the initial position of the intonational phrase ${}_l$, where they cannot attach to a preceding prosodic host: (${}_l = \sigma$). This issue is resolved by positing that p-structure is organised according to its own principles and constraints. One of these constraints is prosodic inversion (Halpern 1995), which allows for the clitics to be placed after the first valid prosodic host – in this case (${}_l / \text{taj} / \omega$). As a consequence, the initial linear order of the phonological string as depicted in Figure 19 will be adjusted to satisfy the prosodic constraints. The result is the (final) p-string ($\text{taj} / \omega = \text{joj} = \text{ga} = \text{je} \text{ (t} \underline{\text{č}} \text{ovek)} / \omega \dots$ which in turn forms the basis for the phonetic representation.

This approach to the interface allows for a clear separation of syntactic and prosodic analyses and can account for a number of other phenomena as well,

²⁹Which approach is chosen for the mapping between syntactic and prosodic constituency is up to the researcher. In this case, the end-based approach would not lead to a different outcome with respect to the clitic placement.

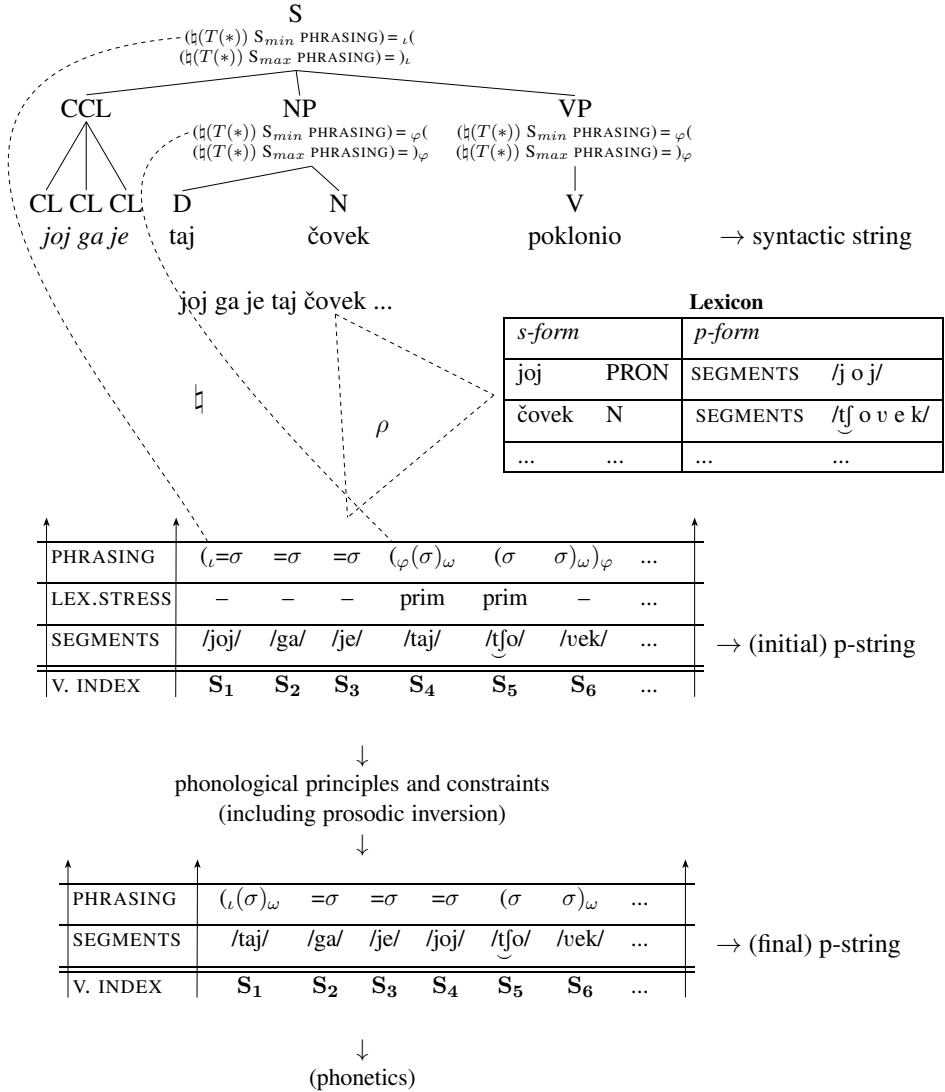


Figure 19: The syntax-prosody interface and the p-structure analysis for example (21)

including notorious problems where the concept of lexical integrity (Bresnan & Mchombo 1995) is seemingly at stake. Such phenomena, among them Pashto enclitics (Bögel 2015: Ch. 6) and Vafsi mesoclitics (Bögel et al. 2018), are difficult to explain from a purely syntactic perspective, but can be explained in a straightforward fashion with the proposal made here, as prosodic restructuring is based on prosodic constraints alone and is not concerned with syntactic constraints like word integrity.

The 2P clitic analysis just discussed was an example for production, i.e., the analysis first considers syntactic structure and then builds prosodic structure. The framework proposed in Bögel (2015) also allows for comprehension as is demonstrated in the following with an example from Butt et al. (2017, 2020) on Urdu polar *kya*. Consider example (23), where the sentence can be understood either as a polar question or as a wh-constituent question.

- (23) *alina=ne zain=ko kya tohfa di-ya t^h-a?*
 Alina=Erg Zain=Acc what present.M.Sg give-Perf.M.Sg be.Past-M.Sg
 Constituent Question: ‘What gift did Alina give to Zain?’
 Polar Question: ‘Did Alina give a gift to Zain?’

This ambiguity corresponds to two different possible syntactic analyses, as shown in Figure 20. In the wh-constituent interpretation, *kya* is phrased together with the following noun *tohfa*. In contrast, in the polar interpretation, *kya* is analyzed as an immediate daughter of S.

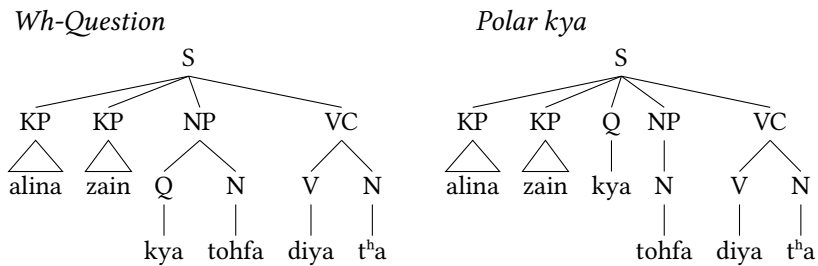


Figure 20: C-structures for the *wh*-reading and for *polar kya*

Prosody is essential to the disambiguation of this structure: *kya* carries an H* accent if it is part of a constituent question while it has a flat pitch in the polar interpretation. In order for the grammar to make use of this disambiguation possibility, the information on pitch accents thus has to be available at the interface to p-structure.

The categorical interpretation of pitch accents is dependent on a number of attributes in a given speech signal; for a H* pitch accent, there needs to be a sudden rise followed by a relatively abrupt drop in the fundamental frequency F_0 . The p-diagram in Figure 21 allows for the integration of this (and additional) speech signal information on the ‘signal’ level (here: medium F_0 and duration for each syllable) with a categorical interpretation of the relevant acoustic cues given on the ‘interpretation’ level in form of a ToBI annotation: H*.³⁰

↑	↑													↑	
PHRAS.	(...)	INTERPRET.
ToBI	H*	L%	↓	
DUR.	0,08	0,16	0,14	0,17	0,28	0,23	0,21	0,20	0,16	0,13	0,11	0,22		SIGNAL	
F_0	164	211	239	243	228	229	247	229	162	147	136	(83)		↓	
VALUE	[ə]	[li]	[na]	[ne]	[zæɪn]	[ko]	[kja]	[toh]	[fa]	[di]	[ja]	[tʰa]			
INDEX	S ₁	S ₂	S ₃	S ₄	S ₅	S ₆	S ₇	S ₈	S ₉	S ₁₀	S ₁₁	S ₁₂			

Figure 21: P-diagram for the speech signal corresponding to the constituent question in (23)

This effectively provides for an interface between phonetics (i.e., a concrete speech signal) and phonology/prosody and allows for the integration of information given in a speech signal into an LFG grammar. The relevant information on the H* accent again becomes available at the interface to c-structure via the transfer of structure and the transfer of vocabulary.

As shown in Figure 22, the transfer of vocabulary identifies the correct p-forms in the multidimensional lexicon and makes the associated s-forms available to c-structure. The transfer of structure then ‘checks’ whether the syllable associated with the c-structure node Q carries an H* accent – in which case the attribute-value pair [QUESTION-TYPE = constituent] is projected to f-structure. This approach also allows for the disambiguation of syntactically ambiguous structures where the boundary domains are the crucial indicators (Bögel 2020) and has also been applied to more complex pitch accent phenomena (Bögel & Raach 2020).

Summing up, the directional indirect reference approach proposed in Bögel (2015) allows for a differentiation of production and comprehension and has been

³⁰The annotation on the interpretation level is greatly reduced on purpose. There is not yet a fully developed ‘UrduToBI’ or a clear conception of possible prosodic domains and how these are defined (but see Urooj et al. 2019 for discussion) – an interpretation in terms of the English/German annotation system might thus be misleading.

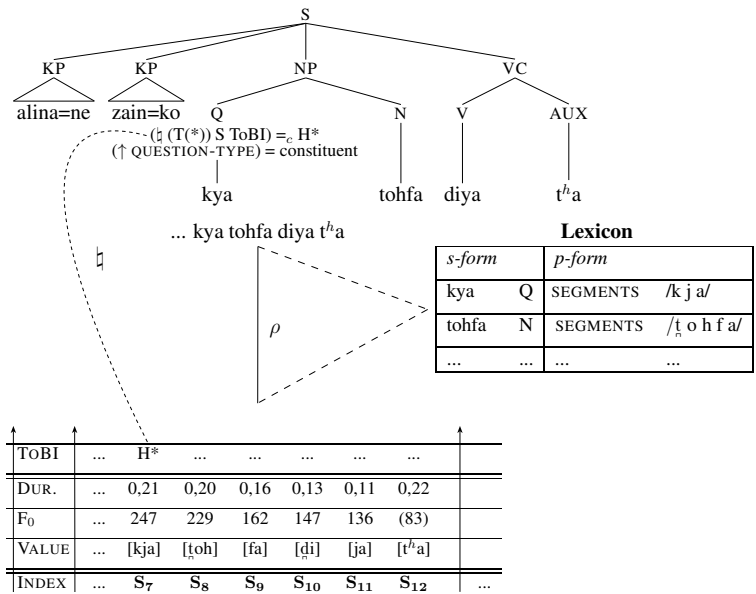


Figure 22: *kya* as part of a constituent question (Butt et al. 2020)

applied to a large variety of different linguistic phenomena. It is the first approach in LFG which integrates spoken language in the form of concrete speech signal data and which pushes LFG towards a more psycholinguistic model of language as discussed in Section 3.

5 Conclusion

This chapter gave a chronological overview of the different approaches to prosody and its interfaces in LFG. As the different proposals show, work at this particular interface always requires a discussion of grammar architecture in general and of module interaction in particular. In general, two schools of thought can be distinguished in the LFG literature: the indirect reference approach and the parallel approach. The indirect reference approach assumes that p-structure is influenced by information from different modules, for example syntactic constituency. In addition, p-structure is assumed to be subject to its own principles and constraints, among them rhythmic principles, prosodic inversion, or constraints on the size of prosodic domains. The indirect reference approach was pursued in Butt & King (1998), O'Connor (2005), and Bögel (2015) and subsequent

papers. While these proposals show differences with respect to the overall architecture, the interfaces between modules in all of these approaches are organised according to the traditional co-descriptive LFG annotations.

The second school of thought, the parallel approach, assumes that modules are built up in parallel. Under this view, each module is built on its own principles and constraints without ‘input’ from the other modules. P-structure is assumed to be formed on rhythmic principles, thus accounting for the mismatches found between prosodic and syntactic constituency. The interface between c- and p-structure is reduced to the interface between the syntactic and the phonological string, which are extended to include prosodic and syntactic/semantic information. The information present in both strings is then ‘matched’. This approach is most prominently pursued in Dalrymple & Mycock (2011), Mycock & Lowe (2013), and subsequent work. Bögel et al. (2009, 2010) also fall into this second group. However, the exact nature of p-structure is never defined under this approach and it is thus harder to demarcate.

A second main point of this chapter was that the majority of the proposals presented here assume a certain directionality, which is also in line with psycholinguistic and computational approaches (Forst & King forthcoming [this volume]): ‘production’ in the case of Butt & King (1998), ‘comprehension’ in Bögel et al. (2009, 2010); O’Connor (2005), and an open discussion of both in Bögel (2015). This distinction is not evident in the proposals by Dalrymple & Mycock (2011) and Mycock & Lowe (2013), which represent a perspective where each module builds structure independently of the other modules. The discrepancy between these views of the grammar architecture and of the analysis of language in general has, to my knowledge, not yet been openly debated. This chapter hopefully contributes to a more general discussion of grammar architectures in that it aims to show in a very concrete way what each school of thought pursues and how these ideas can be realised.

Acknowledgements

I would like to thank Tracy H. King, Miriam Butt, and two anonymous reviewers for their detailed comments, and Mary Dalrymple for her infinite patience.

References

Asudeh, Ash. 2009. Reflexives in the correspondence architecture. Slides of a talk presented at the University of Iceland.

- Baumann, Stefan, Johannes Becker, Martine Grice & Doris Mücke. 2007. Tonal and articulatory marking of focus in German. In *Proceedings of the International Congress of Phonetic Sciences (ICPhS)*.
- Beckman, Mary E., Julia Hirschberg & Stefanie Shattuck-Hufnagel. 2005. The original ToBI system and the evolution of the ToBI framework. In Sun-Ah Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*, chap. 2, 9–54. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199249633.003.0002.
- Beckman, Mary E. & Janet B. Pierrehumbert. 1986. Intonational structure in English and Japanese. *Phonology Yearbook* 3. 255–309.
- Belyaev, Oleg. Forthcoming(a). Core concepts of LFG. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.
- Belyaev, Oleg. Forthcoming(b). Introduction to LFG. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.
- Bennett, Ryan & Emily Elfner. 2019. The syntax-prosody interface. *Annual Review of Linguistics* 5. 151–171. DOI: 10.1146/annurev-linguistics-011718-012503.
- Bögel, Tina. 2015. *The syntax-prosody interface in Lexical Functional Grammar*. Konstanz: University of Konstanz. (Doctoral dissertation).
- Bögel, Tina. 2020. German case ambiguities at the interface: Production and comprehension. In Gerrit Kentner & Joost Kremers (eds.), *Prosody in syntactic encoding* (Linguistische Arbeiten 573), 51–84. Berlin: De Gruyter.
- Bögel, Tina, Miriam Butt, Ronald M. Kaplan, Tracy Holloway King & John T. Maxwell III. 2009. Prosodic phonology in LFG: A new proposal. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '09 conference*, 146–166. Stanford, CA: CSLI Publications.
- Bögel, Tina, Miriam Butt, Ronald M. Kaplan, Tracy Holloway King & John T. Maxwell III. 2010. Second position and the prosody-syntax interface. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '10 conference*, 106–126. Stanford, CA: CSLI Publications.
- Bögel, Tina, Miriam Butt & Sebastian Sulger. 2008. Urdu *Ezafé* and the morphology-syntax interface. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '08 conference*, 129–149. Stanford, CA: CSLI Publications.
- Bögel, Tina & Lea Raach. 2020. Swabian *ed* and *edda*: Negation at the interfaces. In Miriam Butt & Ida Toivonen (eds.), *Proceedings of the LFG '20 conference*. Stanford, CA: CSLI Publications.
- Bögel, Tina, Saeed Reza Yousefi & Mahinnaz Mirdehghan. 2018. Vafsi oblique pronouns: Stress-related placement patterns. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '18 conference*, 88–108. Stanford, CA: CSLI Publications.

- Bresnan, Joan & Sam A. Mchombo. 1995. The lexical integrity principle: Evidence from Bantu. *Natural Language & Linguistic Theory* 13(2). 181–254. DOI: 10.1007/bf00992782.
- Butt, Miriam, Tina Bögel & Farhat Jabeen. 2017. Polar *kya* and the prosody-syntax-pragmatics interface. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '17 conference*. Stanford, CA: CSLI Publications.
- Butt, Miriam, Farhat Jabeen & Tina Bögel. 2020. Ambiguity resolution via the syntax-prosody interface: The case of *kya* ‘what’ in Urdu/Hindi. In Gerrit Kentner & Joost Kremers (eds.), *Prosody in syntactic encoding* (Linguistische Arbeiten 573), 85–118. Berlin: De Gruyter.
- Butt, Miriam & Tracy Holloway King. 1998. Interfacing phonology with LFG. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '98 conference*. Stanford, CA: CSLI Publications.
- Chen, Matthew Y. 1987. The syntax of Xiamen tone sandhi. *Phonology Yearbook* 4. 109–149. DOI: 10.1017/s0952675700000798.
- Dalrymple, Mary. 2001. *Lexical Functional Grammar* (Syntax and Semantics 34). New York: Academic Press. DOI: 10.1163/9781849500104.
- Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell III & Annie Zaenen (eds.). 1995. *Formal issues in Lexical-Functional Grammar*. Stanford, CA: CSLI Publications.
- Dalrymple, Mary, John J. Lowe & Louise Mycock. 2019. *The Oxford reference guide to Lexical Functional Grammar*. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198733300.001.0001.
- Dalrymple, Mary & Louise Mycock. 2011. The prosody-semantics interface. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '11 conference*, 173–193. Stanford, CA: CSLI Publications.
- Elordieta, Gorka. 2008. An overview of theories of the syntax-phonology interface. *Anuario del Seminario de Filología Vasca “Julio de Urquijo”* 42(1). 209–286.
- Féry, Caroline. 2020. Grammatical reflexes of information structure in Germanic languages. In Michael T. Putnam & B. Richard Page (eds.), *Cambridge handbook for Germanic linguistics*, 661–685. Cambridge, UK: Cambridge University Press. DOI: 10.1017/9781108378291.029.
- Forst, Martin & Tracy Holloway King. Forthcoming. Computational implementations and applications. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.
- Frota, Sónia. 2012. Prosodic structure, constituents, and their implementation. In Abigail C. Cohn, Cécile Fougeron & Marie K. Huffman (eds.), *The Oxford handbook of laboratory phonology*, chap. 11, 255–265. Oxford: Oxford University Press.

- Ghini, Mirco. 1993. Phi-formation in Italian: A new proposal. *Toronto Working Papers in Linguistics* 12(2). 41–78.
- Goldsmith, John. 1976. *Autosegmental phonology*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).
- Grice, Martine & Stefan Baumann. 2002. Deutsche Intonation und GToBI. *Linguistische Berichte* 191. 267–298.
- Grice, Martine, D. Robert Ladd & Amalia Arvaniti. 2000. On the place of phrase accents in intonational phonology. *Phonology* 17(2). 143–185. DOI: 10.1017/S0952675700003924.
- Gunlogson, Christine. 2003. *True to form: Rising and falling declaratives as questions in English*. New York: Routledge. DOI: 10.4324/9780203502013.
- Halpern, Aaron. 1995. *On the placement and morphology of clitics*. Stanford, CA: CSLI Publications.
- Hayes, Bruce. 1989. The prosodic hierarchy in meter. In Paul Kiparsky & Gilbert Youmans (eds.), *Rhythm and meter*, 201–260. Orlando: Academic Press. DOI: 10.1016/b978-0-12-409340-9.50013-9.
- Hayes, Bruce & Aditi Lahiri. 1991. Bengali intonational phonology. *Natural Language & Linguistic Theory* 9(1). 47–96. DOI: 10.1007/bf00133326.
- Inkelas, Sharon & Draga Zec. 1995. Syntax-phonology interface. In John A. Goldsmith (ed.), *The handbook of phonological theory*, chap. 15, 535–549. Cambridge, MA: Blackwell.
- Jackendoff, Ray. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.
- Jackendoff, Ray. 2010. The Parallel Architecture and its place in cognitive science. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, chap. 23, 583–605. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199544004.013.0023.
- Jones, Stephen. 2016. The syntax–prosody interface in Korean: Resolving ambiguity in questions. In Doug Arnold, Miriam Butt, Berthold Crysmann, Tracy Holloway King & Stefan Müller (eds.), *Proceedings of the joint 2016 conference on Head-Driven Phrase Structure Grammar and Lexical Functional Grammar*. Stanford, CA: CSLI Publications.
- Kaisse, Ellen M. 1985. *Connected speech: The interaction of syntax and phonology*. Orlando: Academic Press.
- Kaplan, Ronald M. 1987. Three seductions of computational psycholinguistics. In Peter Whitelock, Mary McGee Wood, Harold L. Somers, Rod Johnson & Paul Bennett (eds.), *Linguistic theory and computer applications*, 149–188. London: Academic Press. Reprinted in Dalrymple, Kaplan, Maxwell & Zaenen (1995: 339–367).

- Kaplan, Ronald M. & Martin Kay. 1994. Regular models of phonological rule systems. *Computational Linguistics* 20. 331–478.
- Kingston, John. 2019. The interface between phonetics and phonology. In William F. Katz & Peter F. Assmann (eds.), *The Routledge handbook of phonetics*, 359–400. Abingdon/New York: Routledge. DOI: 10.4324/9780429056253-14.
- Kügler, Frank, Stefan Baumann, Bistra Andreeva, Bettina Braun, Martine Grice, Jana Neitsch, Oliver Niebuhr, Jörg Peters, Christine T. Röhr, Antje Schweitzer & Petra Wagner. 2019. Annotation of German intonation: DIMA compared with other annotation systems. In S. Calhoun, P. Escudero, M. Tabain & P. Warren (eds.), *Proceedings of the International Congress of Phonetic Sciences (ICPhS)*. Australasian Speech Science & Technology Association Inc.
- Ladd, D. Robert. 1986. Intonational phrasing: The case for recursive prosodic structure. *Phonology Yearbook* 3. 311–340. DOI: 10.1017/s0952675700000671.
- Lahiri, Aditi & Jennifer Fitzpatrick-Cole. 1999. Emphatic clitics and focus intonation in Bengali. In René Kager & Wim Zonneveld (eds.), *Phrasal phonology*, 119–144. University of Nijmegen Press.
- Lahiri, Aditi & Frans Plank. 2010. Phonological phrasing in Germanic: The judgement of history, confirmed through experiment. *Transactions of the Philological Society* 108(3). 372–398. DOI: 10.1111/j.1467-968x.2010.01246.x.
- Lehiste, Ilse, Joseph P. Olive & Lynn A. Streeter. 1976. Role of duration in disambiguating syntactically ambiguous sentences. *The Journal of the Acoustical Society of America* 60. 1199–1202. DOI: 10.1121/1.381180.
- Levelt, Willem J. M. 1999. Models of word production. *Trends in Cognitive Sciences* 3(6). 223–232. DOI: 10.1016/s1364-6613(99)01319-4.
- Levelt, Willem J. M., Ardi Roelofs & Antje S. Meyer. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22. 1–75. DOI: 10.1017/s0140525x99001776.
- Liberman, Mark. 1975. *The intonational system of English*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).
- Lowe, John J. 2016. Clitics: Separating syntax and prosody. *Journal of Linguistics* 52(2). 375–419.
- Lowe, John J. & Oleg Belyaev. 2015. Clitic positioning in Ossetic. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '15 conference*. Stanford, CA: CSLI Publications.
- McCarthy, John J. & Alan Prince. 1993. Generalized alignment. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology*, 79–153. Dordrecht: Kluwer Academic Publishers. DOI: 10.1007/978-94-017-3712-8_4.
- McCawley, James D. 1968. *The phonological component of a grammar of Japanese*. The Hague: Mouton.

- Mycock, Louise. 2006. *The typology of wh-questions*. Manchester: University of Manchester. (Doctoral dissertation).
- Mycock, Louise & John J. Lowe. 2013. The prosodic marking of discourse functions. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '13 conference*, 440–460. Stanford, CA: CSLI Publications.
- Nespor, Marina & Irene Vogel. 1986. *Prosodic phonology*. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110977790.
- Nespor, Marina & Irene Vogel. 1989. On clashes and lapses. *Phonology* 6(1). 69–116. DOI: 10.1017/s0952675700000956.
- O'Connor, Robert. 2005. Clitics in LFG: Prosodic structure and phrasal affixation. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG '05 conference*. Stanford, CA: CSLI Publications.
- Pierrehumbert, Janet B. 1980. *The phonology and phonetics of English intonation*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).
- Pierrehumbert, Janet B. & Mary E. Beckman. 1988. *Japanese tone structure*. Cambridge, MA: The MIT Press.
- Prince, Alan & Paul Smolensky. 2004. *Optimality theory: Constraint interaction in generative grammar*. Oxford/Malden: Blackwell.
- Sadock, Jerrold M. 1991. *Autolexical Syntax: A theory of parallel grammatical representations*. Chicago: University of Chicago Press.
- Schütze, Carson. 1994. Serbo-Croatian second position clitic placement and the phonology-syntax interface. In Andrew Carnie, Heidi Harley & T. Bures (eds.), *MIT Working Papers in Linguistics: Papers on phonology and morphology*, 373–473. Cambridge, MA: Department of Linguistics & Philosophy, MIT. Revised version.
- Selkirk, Elisabeth O. 1978. On prosodic structure and its relation to syntactic structure. In Thorstein Fretheim (ed.), *Nordic prosody ii*, 111–140. Tapir.
- Selkirk, Elisabeth O. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: The MIT Press.
- Selkirk, Elisabeth O. 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3. 371–405. DOI: 10.1017/s0952675700000695.
- Selkirk, Elisabeth O. 1995. The prosodic structure of function words. In Jill N. Beckmann, Laura Walsh Dickey & Suzanne Urbanczyk (eds.), *Papers in optimality theory*. University of Massachusetts: GLSA.
- Selkirk, Elisabeth O. 2009. On clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure. *Gengo Kenkyu (Journal of the Linguistic Society of Japan)* 136. 35–73.

- Selkirk, Elisabeth O. 2011. The syntax-phonology interface. In John A. Goldsmith, Jason Riggle & Alan C. L. Yu (eds.), *The handbook of phonological theory*, 435–484. Malden, MA: Blackwell. DOI: 10.1002/9781444343069.ch14.
- Shattuck-Hufnagel, Stefanie & Alice Turk. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25(2). 193–247. DOI: 10.1007/bf01708572.
- Silverman, Kim, Mary Beckman, John Pitrelli, Mari Ostendorf, Colin Wightman, Patti Price, Janet B. Pierrehumbert & Julia Hirschberg. 1992. ToBI: a standard for labeling English prosody. In *Proceedings of the 1992 international conference on spoken language processing*. Banff.
- Sweet, Henry. 1904. *Elementarbuch des gesprochenen Englisch*. 3rd edn. Leipzig: Tauchnitz [u.a.]
- Urooj, Saba, Benazir Mumtaz & Sarmad Hussain. 2019. Urdu intonation. *Journal of South Asian Linguistics* 10. Special issue on the prosody of South Asian languages.
- Varaschin, Giuseppe. Forthcoming. LFG and Simpler Syntax. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.
- Wagner, Michael. 2010. Prosody and recursion in coordinate structures and beyond. *Natural Language & Linguistic Theory* 28. 183–237. DOI: 10.1007/s11049-009-9086-0.
- Wells, John C. 1970. Local accents in England and Wales. *Journal of Linguistics* 6(2). 231–252. DOI: 10.1017/s0022226700002632.
- Zaenen, Annie. Forthcoming. Information structure. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*. Berlin: Language Science Press.
- Zaenen, Annie & Ronald M. Kaplan. 1995. Formal devices for linguistic generalizations: West Germanic word order in LFG. In Jennifer S. Cole, Georgia M. Green & Jerry L. Morgan (eds.), *Linguistics and computation*, 3–27. Stanford, CA: CSLI Publications. Reprinted in Dalrymple, Kaplan, Maxwell & Zaenen (1995: 215–240).