Prosodic phrasing and the emergence of phrase structure

Nikolaus P. Himmelmann
Universität zu Köln

REVISED 22 APRIL 2021

Abstract

In order to clarify the role of prosodic phrasing in the emergence of phrase structure, it is necessary to be clear as to how syntactic phrasing relates to prosodic phrasing. The core proposal here is that a distinction has to be made between two basic types of syntactic constructions; namely, syntactic configurations for which prosodic phrasing is part of the definition of the construction, and phrase structure proper, which is essentially independent of prosodic phrasing. This distinction implies a somewhat more complex view of the interface between syntax and prosody than is currently widely assumed. Rather than conceiving of prosodic phrases as derivative of syntactic phrases, prosodic phrasing and syntactic phrase structure are seen here as alternative ways of relating words to each other, forming larger units from smaller ones. Against this background, the emergence of phrase structure is conceived of as the emancipation of syntax from prosodic scaffolding.

Keywords: intonation unit, constituent structure, syntax-prosody interface, right detachments, clitic doubling

1. Introduction

This paper is concerned with the level of prosodic phrasing widely known as *intonation unit* (IU) and its relation to syntactic phrasing. It is proposed that a distinction needs to be made between prosodically robust phrase structure and syntactic configurations where prosody plays a key role in determining meaning composition. The latter will be referred to as *prosody-dependent constructions*. They can be of two types. One type is called *prosodic groupings* and includes constructions involving two or more prosodic phrases such as afterthought constructions and (loose) apposition. The other type comprises *IU-bounded constructions* for which it is a definitional requirement that all constituents occur within a single IU (e.g. serial verb constructions). Only prosodically robust constructions are constituent structures in the proper sense and thus are amenable to typical constituent structure tests such as being movable as a single unit. Phrase structures in this sense are the result of diachronic processes and thus may arise through grammaticization, for example.

The argument is structured as follows. Apart from providing an operational definition of IUs, Section 2 briefly reviews the literature on the syntax-prosody interface, focusing on work based on the investigation of non-scripted (rather than laboratory) speech. The main finding relevant to the present context is that there are very few, if any, hard constraints on the *size* of syntactic structures found in IUs. IUs may consist of as little as a single syllable and extend across two or more sentences, the upper limit being determined by processing constraints, including the need to breathe. However, there are constraints, or at least strong tendencies, with regard to the alignment of the boundaries of syntactic and prosodic phrases. With some principled exceptions, these boundaries tend to overlap.

Section 3 starts with the observation that for standardly assumed phrase structures such as prepositional phrases (PPs) or determiner phrases (DPs)[1], the prosodic packaging is never a concern and is never mentioned as part of their definition. Even a very heavy DP with multiple layers of modification is a DP regardless of whether it is produced in a single IU or spread across a number of IUs. Phrase structures of this type are thus prosodically robust, i.e., independent of their prosodic packaging. In this regard, they differ from other types of commonly recognized syntactic configurations such as left- or right-detached constructions (topics, afterthoughts, etc.), parentheticals, or serial verbs for which prosodic packaging is often considered to be a constitutive property of the construction. Inasmuch as prosodic packaging is constitutive of these latter constructions, they may be aptly characterized as prosody-dependent constructions.

In section 4, it is noted that the importance of prosodic phrasing for the development of phrase structures is probably somewhat limited: the grammaticization of prosodically robust phrase structures presupposes that the constituents of an emerging phrase regularly occur within a single IU. This is a necessary but not a sufficient condition for the further grammaticization of phrase structure. It is likely that this condition holds for many candidate structures in all languages at all times, including for example all of the different types of nominal expressions distinguished in section 3.1 of the introduction to this special issue. But not all languages have developed prosodically robust phrase structures. Hence, seemingly paradoxically, the grammaticization process leads to the autonomy of syntax from prosody: fully grammaticized phrase structures are independent of prosodic scaffolding, i.e. are prosodically robust.

Section 5 summarizes.

At present, it is unclear how exactly the distinction between prosodically robust phrase structure and prosody-dependent constructions maps onto the different construal types for nominal expressions introduced in section 3 of the introduction to this special issue. For the purposes of the present argument, 'phrase structure proper' is equivalent to 'rigidly structured' phrase structure as defined there. Whether and to what extent the other construal types proposed in the introduction also belong to the prosodically robust phrase structures needs further investigation. Some further remarks on this topic follow in section 4 below.

It should be noted right at the outset that even if I occasionally refer simply to prosody, this article is strictly limited to prosodic phrasing. It is very well possible that the interaction between syntax and prosodic prominence is somewhat differently organized. More specifically, the claim has repeatedly been made that prosodic prominence ('stress' or 'accent') plays a major role in the historical development of syntactic structures (e.g. Vincent 1999 with reference to the development of PPs in Romance languages). Furthermore, constructions that involve grammatical tone are also of no concern here. This, for example,

---

[1] In referring to standardly assumed phrase structures, I will make use of widely used terminology and abbreviations such as DP, PP, VP and CP, unless reporting from the literature. This does not mean that I would subscribe to all assumptions that are often associated with this terminology. Specifically, I would not want to claim that all nominal expressions are determiner phrases by default (rather, a DP requires an overt determiner position) or that all clauses are CPs by default. Cp. Pullum 1985, McCawley 1989, Himmelmann 1997, Matthews 2007, Reinöhl 2016, Börjars et al. 2016, Bruening 2020, inter alia, for pertinent discussion.
As in most other cross-linguistic work, PP here includes postpositional phrases, hence *adpositional phrases* would be the more precise term. Finally, I use the functionally defined terms *nominal expression, adpositional expression* and *event expression* when it is necessary to be non-specific about the precise syntactic structure of a particular group of words (i.e. whether it is a hierarchically organized phrase or a loosely adjoined group).

includes cases where the difference between a main and a subordinate clause or a future and a past tense is conveyed by a specific tonal pattern.

2. Intonation units and syntactic units

Spoken language is produced in melodically and rhythmically coherent chunks. Here, the term *intonation unit* (IU) is used in reference to these chunks. A widely used alternative is *intonational phrase* (IP). The former is the term employed by Chafe (1994) and colleagues, whose work is based on non-scripted data, including conversational interactions.

While definitions for IUs are somewhat variable in their scope (cp. Ladd 2008:288-299; Himmelmann et al. 2018:210-214), it is clear that, especially in the context of an investigation of the syntax-prosody interface, the definition of prosodic units must not refer to syntactic units (unless one starts with the assumption made, for example, in Prosodic Phonology that prosodic units are derivable from syntactic units, as further discussed in section 3). Instead, prosodic units need to be defined with reference to prosodic features, which in turn have to be phonetically explicit in order to be operationalizable (Ladd 2008:289).

The empirical validity of the following definition of an IU has been thoroughly tested in an interrater agreement study by Himmelmann et al. (2018). There are two basic features that lead to an interruption of the intonational coherence of a chunk of speech and thus define the boundary between two IUs (see also DuBois et al. 1992):

1. An interruption of the rhythmic delivery by a pause, lengthening of the last segment at the end of a unit and/or increased speaking rate at the beginning of a new unit.
2. A disruption of the pitch contour/melody line: a pitch jump (up or down) between the end of one unit and the beginning of the next.

Strictly speaking, only the second, melodic criterion is essential, but this is often rather difficult to identify. The rhythmic criteria are optional, but when they are present, they significantly contribute to the strength and, concomitantly, the perceptual clearness of an IU boundary. Boundaries involving a pause are particularly easy to perceive, but boundary-marking pauses must be distinguished from IU-internal hesitation pauses. These criteria identify IUs as *phonetic* units which, as argued in Himmelmann et al. (2018), are crosslinguistically identifiable in this way. In many languages, IUs receive additional language-specific marking, for example through the use of boundary tones. Example (1) and Figure 1 illustrate a simple, clear IU boundary in a German spontaneous narrative.[2]

(1) un jetz kommt's (0.4)
    and now   comes=it
    das bittere ende (0.5)
    the bitter   end
    'And here/now it comes. The bitter ending.'   [Kölnkorpus FrauHolle_191]

---

[2] The following conventions are used in examples of spontaneous speech: each line is one IU; (=) indicates a boundary without a perceptible pause (sometimes called *latching*); pause length is given between parentheses; < > enclose false starts. No capitals and no punctuation are used in these examples to indicate that this is not standard written language. Sources for the example are given in [ ]. Some examples are from the Kölnkorpus, a collection of spontaneous narrative recordings contributed by students as part of a recording and transcription exercise.

Figure 1: Waveform and F0 track for the German example in (1).
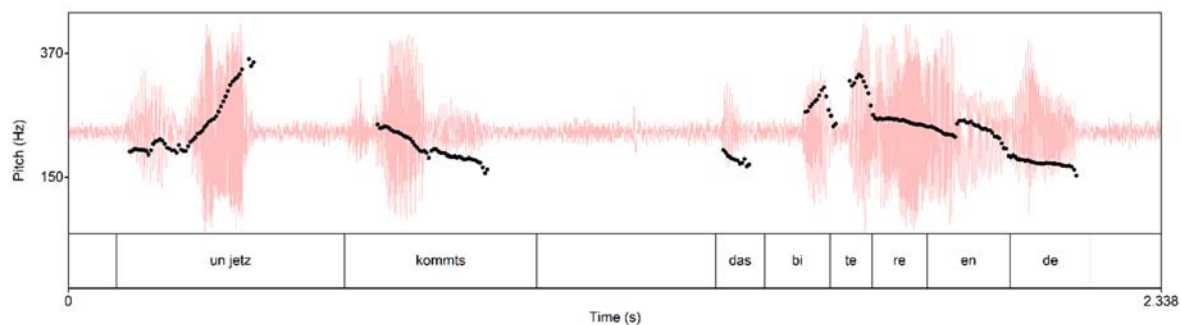
Both IUs in (1) end on a falling pitch contour and are followed by a pause. Onset F0 in the second IU (*das bittere ende*) is higher than offset F0 in the first IU, i.e. there is a pitch jump upward from the first to the second IU. Note that F0 tracks are interrupted whenever there are voiceless consonants but pitch tends to rise or fall continuously across such interruptions. Hence, the gap between *jetz* and *kommts* in the first IU is heard as a continuous fall, not as a pause and a pitch reset. Finally, fully independent consecutive IUs tend to reach similar F0 maxima as also illustrated by this example.

With regard to the syntactic units commonly found in intonation units, Chafe (1994:63f) proposes a fundamental distinction between regulatory and substantive units. Regulatory IUs tend to be short (on average not more than one phonological word long) and regulate interaction and information flow. Typical examples include interjections and short phrases such as *and then*, *yeah, well*, *you know*, etc. Substantive units convey states of affairs or make reference to entities, etc. They are typically 3–5 phonological words long (Chafe 1994:65, see also Croft 1995:873, who provides additional references). The figures vary a little across languages (e.g. Tao 1996:52-54, Ross et al. 2016), but overall there is a strong tendency for IUs in natural speech to be relatively short, especially when compared to the units delimited by commas and full stops in writing.

The preceding observations clearly show that the size of IUs in terms of segments and syntactic units is highly variable, which does not mean, however, that the syntactic content of IUs is completely arbitrary. There are two kinds of constraints. On the one hand, clear tendencies exist regarding the syntactic units that typically occur in an IU. These have been observed for spontaneous speech across a number of different languages. One such tendency is that roughly 50% of IUs in a particular corpus tend to be simple clauses, typically consisting of a verbal predicate and at most two further phrases (DPs or PPs). Table 1 reproduces the figures found in Croft (2007).

Table 1:   Clausal IUs in corpora from five languages (= Table 3 in Croft 2007:12; with one minor modification).

	Clausal IUs (%)	Total IUs (n)	Source
English	47.8%	1,989	Croft 1995:849
Wardaman	50.5%	2,072	Croft 2007:12
Mandarin Chinese	47.9%	1,675	Tao 1996:72
Japanese	57.0%	1,600	Matsumoto 2000:531, Table 1
Korean	55.8%	n/a	Park 2002:642, citing Kim 1996

Among the remaining 50%, noun phrases and prepositional phrases typically belong to the most frequent subtypes, but here there is considerably more variation across languages. Thus, Croft (1995:845) reports 13.7% for "phrases" in English, including both NPs and PPs; Tao (1996:72) has 25.9% NPs and 2% PPs for Mandarin Chinese; and Croft (2007:11) reports 21.1% NPs and no PPs for Wardaman. Other categories are more difficult to compare across different studies, as the authors make use of different categories. Croft (1995:845, 2007:11), for example, refers to "lexical IUs" (i.e. single word IUs), "relative clauses", "coordinate sentences", "triple coordinate sentences", and so on. Tao (1996:72) lists "discourse marker", "pause filler", "adverb", "attributive adjective", inter alia. In sum, there is a clear, cross-linguistically robust tendency for IUs to consist of a single clause or phrase, but a substantial number of IUs are smaller (single words) or larger (multiple clauses) than a clause.

The second, somewhat stronger constraint pertains to the fact that the boundaries of an IU are typically aligned with the boundaries of a syntactic unit. Thus, for example, *the yellow book on the table* is typically included in a single IU or it is chunked into two IUs [*the yellow book*]IU [*on the table*]IU. What occurs only rarely is a chunking that completely ignores syntactic boundaries as in [*the yellow book on*]IU [*the table*]IU or [*the yellow*]IU [*book on the table*]IU. Croft (1995:494f) proposes the *Full Grammatical Unit Condition* to capture this constraint: an IU typically includes a complete syntactic unit, i.e. a clause with the full set of verbal arguments, a determiner phrase with all modifiers and complements, etc. Exceptions to this generalization tend to be systematic. A typical example for a systematic exception is the tendency to distribute overly heavy constituents across two or more IUs (e.g. *the yellow book on the table which we were given on the occasion of our fiftieth wedding anniversary*). More exceptions will be discussed in section 3.

As an interim summary, we may say that the relation between IUs and syntactic units seems to be both flexible and constrained at the same time. It is flexible in that there do not appear to be rigorous rules such as "clauses have to be delivered in a single IU" or, conversely, "every IU consists of exactly one clause". But the relation is constrained in that the boundaries of IUs and syntactic units typically align. And there are strong tendencies; for example, that clauses typically occur within a single IU and that many IUs consist of a single clause.

Within Prosodic Phonology (Selkirk 1984, Nespor & Vogel 1986), various attempts have been made to further formalize the constraints in the relation between prosodic und syntactic units. A core concept in this approach is the idea that there is a hierarchy of prosodic units that

roughly corresponds to the hierarchy of syntactic units, at least on the higher levels. This is illustrated in (2), where *intonational phrase* corresponds to IU as used here.

(2) Intonational Phrase (ɩ)   clause (CP)
     Phonological Phrase (φ)   phrase (DP, PP)
     Prosodic Word (ω)   morphosyntactic word
     Foot
     Syllable

In more recent work, which takes Optimality Theory as the basic theoretical framework, different types of alignment constraints have been proposed to capture the correspondence between prosodic and syntactic unit boundaries (e.g. Truckenbrodt's (1999) *Wrap Theory* or Selkirk's (2011) *Match Theory*). This work has been successful in identifying and providing solutions for a number of instances in which prosodic and syntactic boundaries do not properly align, as further discussed in the next section. But it also tends to significantly underestimate the flexibility in the correspondence between prosodic and syntactic units as far as the overall size of the units is concerned. As just noted, it is not the case that all clauses are mapped onto IUs, and not all IUs contain clauses. This flexibility is better captured by Chafe's (1994) proposal that IUs are essentially units of information, the most basic constraint being Chafe's "one new idea" constraint (Chafe 1994:108 passim; cp. also Pawley & Syder 2000 for a related proposal, Himmelmann et al. 2018:236f for further discussion and references and Dehé 2014:108-110 for a discussion of Selkirk's (1984) *sense unit condition*).

A second major difference between approaches founded on the basic assumptions of Prosodic Phonology and the approach advocated here pertains to what could be called the uniformity assumption. According to the uniformity assumption, which underlies most work on Prosodic Phonology, the relation between syntactic and prosodic units is always of the same type (it essentially involves matching units of either type with each other). Here, it is proposed that the relation may be of two basic types which involve different generalizations, as further detailed in the following section.[3]

3. Prosodically robust phrase structure vs. prosody-dependent constructions

When theorizing the correspondence between syntactic and prosodic structure, it is common practice to look at constituent structures which are rigidly structured, specifically DPs, PPs, VPs and CPs, and more rarely at adjectival and adverbial phrases. Importantly, rigidly

---

[3] This is not the place for an in-depth discussion of different conceptions of the syntax-prosody interface. Scheer (2010) presents a detailed and voluminous historical overview of different theories in the tradition of American structuralism. Elfner (2018) provides a short, slightly updated summary. Of particular interest in the current context is recent work by Richards (2016) and, especially, Wagner (2010, 2015). Richards (2016) proposes a framework which allows direct phonological input into the overall syntactic derivation, an option rigorously excluded in strictly modular approaches. Wagner (2010) develops a new approach to standard examples of putative syntax-prosody mismatches (cp., for example, example (5) below). While adhering to a unidirectional view (syntactic units are mapped onto prosodic ones and not the other way around) and also to the uniformity assumption, he proposes a reanalysis of the syntactic structure in such a way that both the syntax and the prosody of such examples become more list-like (or flatter), which in part overlaps with the view developed in the present paper. Wagner also highlights the importance of relative prosodic boundary strength, allowing for recursive embedding of prosodic units, as opposed to categorically different boundary types, another point of similarity with the present approach.
All these works, unfortunately, lack references to other traditions, most importantly the work inspired by Bolinger, Halliday, Crystal and Chafe (see Schultze-Berndt 2002:276f and Ladd 2008). The issue is also a major concern in conversation analysis (e.g. Couper-Kuhlen 1986; Auer 1991, 1996; Selting 1995).

structured phrase structures all have in common the fact that they are defined completely independently of prosodic considerations. In this regard, they differ from another, more heterogeneous set of syntactic configurations such as left-detached topics, right-detached afterthoughts, or parentheses, where prosody is deemed to be a constitutive part of the construction. Somewhat surprisingly, these constructions are rarely, if ever, mentioned in discussions pertaining to the syntax-prosody interface. The claim here is that this is not simply an oversight. Rather, these two types of syntactic constructions clearly differ in their relation to prosodic phrasing.

3.1 Prosodically robust phrase structure

One basic property that characterizes rigidly structured phrase structures is the fact that they function as units within a larger domain. The well-known constituent structure tests, such as permutation or substitution tests, exploit this property by providing evidence that a string of two or more words behaves as a unit with respect to adjacent words that do not belong to the phrase. This single-unit behavior obtains regardless of the prosodic packaging. A heavy DP such as *my uncle's new boat we talked about last night* will usually be split into two IUs (i.e. [*my uncle's new boat*]IU [*we talked about last night*]IU) when used in everyday conversation, but that does not change its syntactic structure. That is, prototypical constituent structures are recognizable and diagnosable as such regardless of their prosodic packaging. Example (3) represents an extreme case that shows that even the widely met requirement that prosodic boundaries match syntactic boundaries can occasionally be dispensed with. In (3) this requirement is not met for the complementizer *dass* in (3)b and the remainder of the complement clause introduced by it, and the combination of preposition and definite article *von der* and its nominal complement *strecke* in (3)c.

(3) a. un: dann hab ich plötzlich von weitem (0.5)
     and then have I suddenly from afar

   b. gesehen **dass** (0.8)
     seen COMP

   c. en teil **von der** (1.0) äh öh (0.5)
     a part of the FILLER

   d. strecke (=)
     route

   e. öh mit schnee öh (0.7) ähm (1.0)
     FILLER with snow FILLER FILLER

   f. also (0.3)
     well

   g. mit schnee bedeckt war (0.5)
     with snow covered was   [Kölnkorpus Seltsam_17ff]

Figure 2: Waveform and F0 track for the German example in (3).
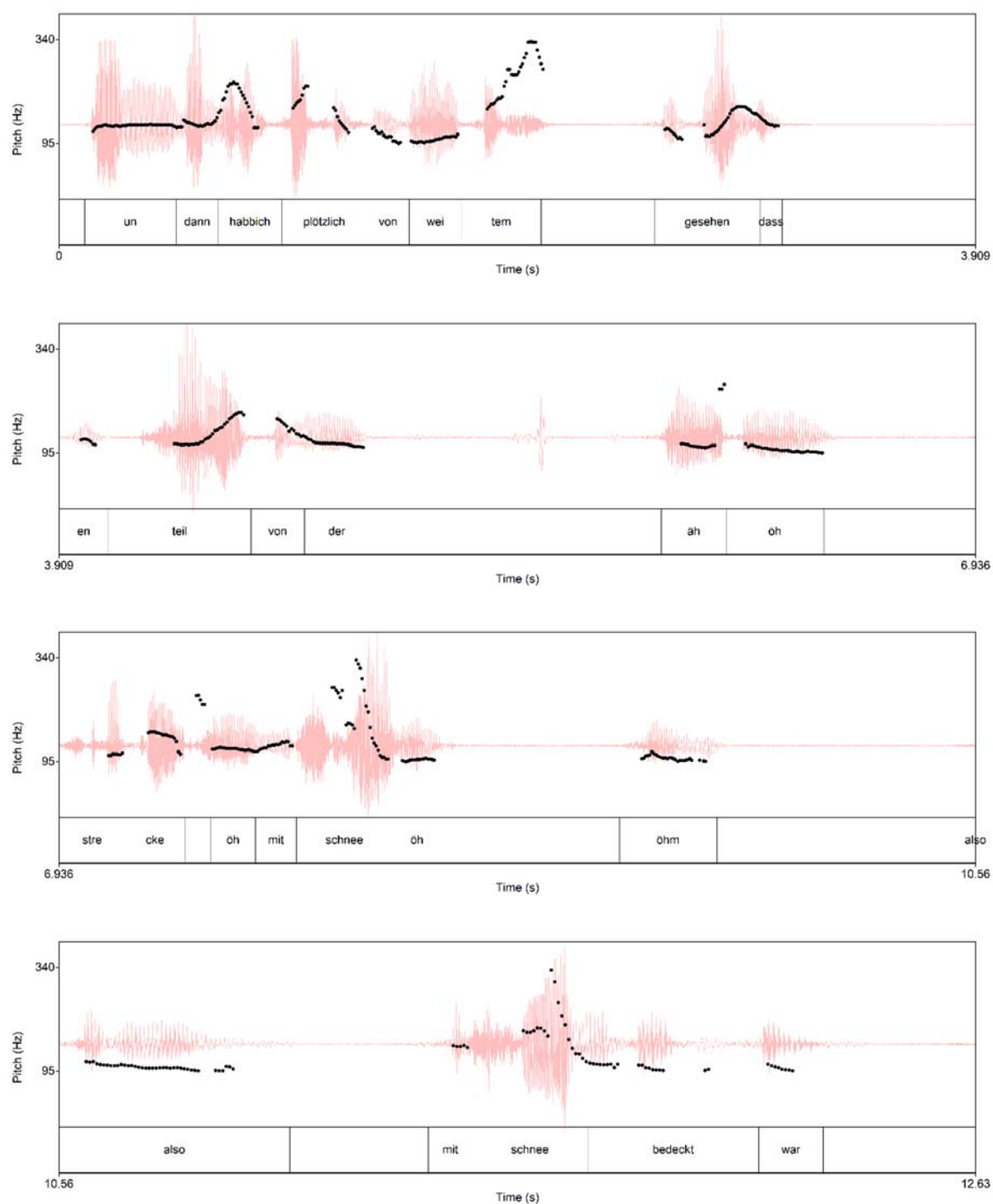
In example (3), the complement clause that functions as the object of the verb *see* in the matrix clause is produced in a fragmentary fashion, split across 5 IUs (b–e, g), interspersed with a regulatory unit (f) and two long filled pauses (in c and e). The "clean" version of this example is given in (4).

(4) un: dann hab ich plötzlich von   weitem gesehen
     and then have I   suddenly   from   afar   seen
     [dass ein teil von der   strecke mit   schnee bedeckt   war]$_{CP}$
     COMP a   part of   the   route   with snow   covered   was
     'And then, I suddenly saw from afar that a part of the route was covered with snow.'

The important point to note here is the fact that despite the extremely fragmented delivery, the complement clause in (3) is recognizably a German complement clause which is fully grammatical, showing the required word order (subject precedes adjunct, verb is final), subject-verb agreement and proper case marking. Similarly, the complex DP [$_{DP}$ *ein teil* [$_{PP}$ *von* [$_{DP}$ *der strecke*]]] is clearly recognizable and fully grammatical despite the very long filled pause separating the final noun from the rest. That is, German prototypical phrase structures such as DPs and CPs are *prosodically robust* and in this regard differ from syntactic constructions which depend on their prosodic packaging.

A particularly important property of prosodically robust phrase structures is the fact that they do not necessarily require precise alignment with prosodic boundaries. In (3)b, for example, the complementizer *dass* forms a prosodic constituent with the matrix verb and is thus separated from the constituent it belongs to. Similarly, in (3)c–d, preposition and determiner are separated from the complement noun *strecke* in the PP *von der strecke*, and thus fail to meet either Croft's *Full Grammatical Unit Condition* or the substantially equivalent alignment constraints in OT-based accounts. While misalignments of function words happen only under special conditions in English and German, in other languages function words such as complementizers and determiners systematically occur in the prosodic unit preceding the one containing the phrase they belong to, as discussed and exemplified more fully in Himmelmann (2014).

The systematic lack of an exact correspondence between prosodic and syntactic constituency in the case of English object relative clauses has in fact been one of the major stepping stones in the development of Prosodic Phonology. Chomsky & Halle (1968:372) famously note that the prosodic chunking of a relative clause modifying the object argument systematically fails to align with the syntactic structure in that it separates the relative clause from its head, as seen in (5), where // indicates an IU boundary.

(5) a.   [$_{CP}$ This is [$_{DP}$ the cat [$_{CP}$ that caught [$_{DP}$ the rat [$_{CP}$ that stole the cheese.]]]]]
     b.   This is the cat // that caught the rat // that stole the cheese.

Hence it is clear that the correspondence between prosodic and syntactic boundaries is also somewhat flexible (see Wagner 2010:224-228 for an alternative analysis).[4]

3.2 Prosodic groupings

Turning now to the second type of syntactic construction, that which lacks prosodic robustness, we find a broad range of constructions which all have in common the fact that prosody is mentioned as a constitutive part of their definition. This includes, for example,

---

[4] Note that prosodic robustness is a characteristic of phrase structures proper regardless of whether or not they are hierarchically organized. For (modern) European languages, phrase structures typically require that constituents occur in an adjacent and largely fixed order, and that they are hierarchically organized, with each phrasing level arguably having a head. But these two properties do not have to go hand in hand, as shown by Tagalog, where nominal and adpositional expressions are rendered by phrasally organized structures which clearly behave as units, but do not exhibit an elaborate hierarchical structure (see Himmelmann 2015 for examples and details). See also the literature referenced in footnote 1.

right-detached constructions such as French *Jean la voit souvent, Marie* 'John often sees her, Mary'. In discussing this example, Kayne remarks:

> There is an intonation contour specific to dislocation constructions (…) that is indicated by the comma placed before the dislocated phrase *Marie*. (Kayne 1994:79)

If there were no such specific intonation contour, it would be a different type of construction, i.e. a clitic doubling construction, which is ungrammatical in French (***Jean la voit souvent Marie*), unlike in Spanish.[5] Similarly, Michaelis & Lambrecht (1996) propose a distinction between Right Dislocation (RD) as in *They're red LEATHER, the shoes she's wearing* and Nominal Extraposition (NE) as in *It's AMAZING the people you SEE here*. The distinction essentially rests on prosodic differences:

> In RD, the postpredicate NP has a low and flat intonation contour, indicating that it follows the right boundary of the VP focus domain. In NE, by contrast, this NP is necessarily accented. (Michaelis & Lambrecht 1996:223)

Other constructions for which prosodic evidence is mentioned in their definition include left-detached constructions (topics), parentheses, (loose) appositions, and quoted speech. Clearly, this is a heterogeneous collection of constructions and no claim is made here that they have anything more in common than the fact that prosody plays a role in their definition.[6] For current purposes, the main point is that they all are not prosodically robust, because the construction changes when the prosodic packaging changes: a right-detached afterthought is no longer a right-detached afterthought if no prosodic boundary occurs between it and the preceding clause. That is, the overall construction consisting of the preceding clause and the right-detached constituent is prosody-dependent. Because they extend beyond an IU boundary, constructions of this type are called *prosodic groupings* in this article. As further discussed in section 4, prosodic groupings constitute one basic type of prosody-dependent constructions. The other basic type involves IU-bounded constructions.

Prosodic groupings are defined by the combination of a morphosyntactic and a prosodic configuration. The basic properties of prosodic groupings are illustrated below with one type of right-detached construction, the so-called clitic right dislocations.[7] Morphosyntactically, clitic right dislocations are defined by the occurrence within the host clause of a pronoun that is co-referential with the right-detached nominal expression. In (6), the pronoun *die* occurs in

---

[5] Somewhat parodoxically, Kayne (1994) actually argues that *Jean la voit souvent, Marie* underlyingly involves clitic doubling, i.e. that it is derived from the ungrammatical **Jean la voit souvent Marie*. See de Cat (2007) for counterarguments. The construction with a prosodic break, i.e. *Jean la voit souvent, Marie*, is also known as *clitic right dislocation* (CLRD), which primarily differs from the clitic doubling construction with regard to the prosodic break. We will return to this point in section 4.

[6] All of these labels tend to subsume a broad range of constructions and it is often difficult to make generalizations that actually hold for all constructions that have been attributed the label at hand. For the purposes of this article, all claims are basically made in reference to what may be considered prototypical instances of a given construction. Thus, *apposition*, for example, refers to *loose apposition*. For the prosody of English parentheses and appositions, Dehé (2014) provides a thorough, corpus-based investigation. It may be useful to note in passing that apposition is primarily found in what Dehé (2014:120) calls "more formal registers of spoken language", substantially more than half of her data actually coming from scripted monologues, in particular news items and academic speeches.

[7] Detachment constructions, both to the left and to the right, have received considerable attention over the last two decades, in different frameworks and on the basis of diverging data sets. It is impossible in the present context to attempt even a cursory overview of the literature. Note, however, that German right-detached constructions here only serve as an example for prosodic groupings. No new claims regarding their form or function are presented. For more discussion of the specifics of right detachments in German, see Altmann (1981) and Averintseva-Klisch (2009), inter alia.

object function and is co-referential with the DP *die schoh*; in (7) the pronoun *die* also occurs in object function and is co-referential with *die samen*. In both examples, the non-finite (participial) part of the verbal complex (*usjetrocke* in (6) and *abgemacht* in (7)) clearly marks the right syntactic boundary of the host clause.

(6) **die**   hät dä mir immer   usjetrocke   (0.1)
     them has he   me always taken.off

     die   **schoh** (0.7)
     the   shoes
     'He always took them off me, the shoes.'   [Bhatt & Lindlar 1998:80 / 2_01dieSchoh]

(7) ich hab **die**   **alle** abgemacht **die**   **samen** (0.5)
     I   have them all   removed   the   seeds
   'I removed them all the seeds.'   [Kölnkorpus Oma_089]

As for the prosodic packaging, there are different possibilities. In (6), there is a clear prosodic boundary between the host clause and the detached DP, as seen in Figure 3. This break consists of a boundary tone on the final unstressed syllable of *usjetrocke*, a short pause, and a major pitch jump downward, making it clear that *die schoh* occurs in an IU of its own. However, as the pitch peak on *schoh* is considerably lower than the one on *immer*, this IU is also clearly heard as an add-on to the preceding one, rather than as a sequence of two IUs of equal standing.

Figure 3: Waveform and F0 track for the German example in (6).
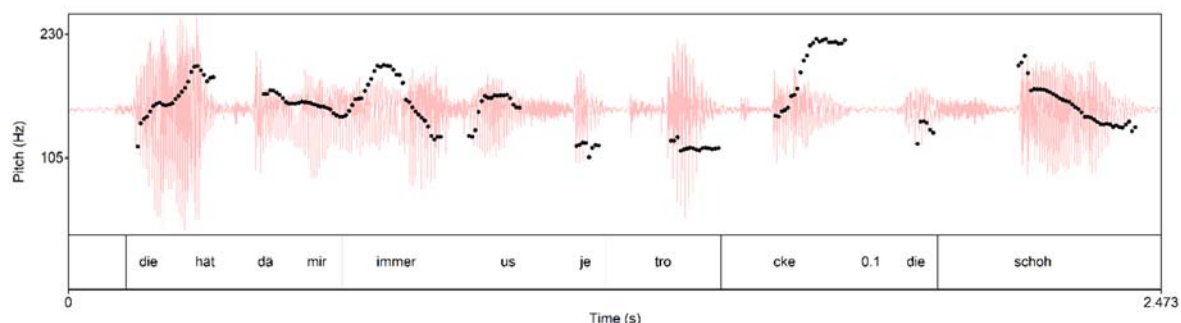
In example (7), the final DP *die samen* does not form an IU of its own but is integrated into the preceding IU. However, the lowest pitch level in the IU has already been reached on *abgemacht* and thus prior to the final DP (Figure 4). There is no pitch movement at all on the DP, which simply continues on the low level already reached on *abgemacht*. The intensity is also greatly reduced. Thus, while not forming an IU of its own, *die samen* extends beyond an IU boundary and in this sense fits the definition of a prosodic grouping given above.

Figure 4: Waveform and F0 track for the German example in (7).
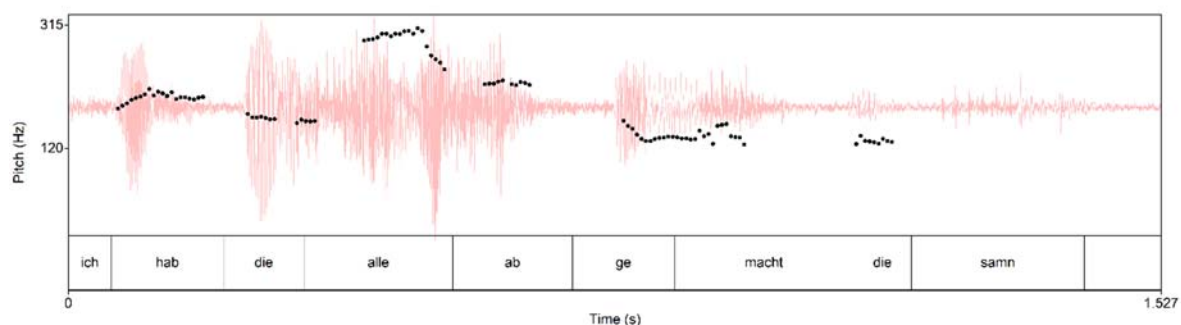
Importantly, in both examples the final DP is prosodically packaged as an add-on in that the endpoint of the IU has already been reached before the DPs are uttered. The two examples differ in how the add-on is prosodically effected. In (6), the DP forms an IU of its own, but a subordinate one. In (7), it is an IU extension (which one could also call a prosodic clitic). Both examples clearly differ from the way the morphosyntactically identical example in (1), repeated in (8) below, is prosodically packaged. As can be seen in Figure 1, in (8) the two IUs are on equal footing in that they basically have the same prosodic structure. They begin with an early high rise and reach very similar heights, after which follows a continuous fall across the remaining syllables towards a final low boundary tone. That is, the second IU, which contains the DP, is not a subordinate add-on to the first one (as in (6)), but rather constitutes a continuation on the same hierarchical level as the first IU.

(8) un jetz kommt's (0.4)
     and now comes=it
     das bittere ende (0.5)
     the bitter end
     'And here/now it comes. The bitter ending.'   [Kölnkorpus FrauHolle_191]

The three different ways in which the final DPs are prosodically packaged in examples (6)–(8) correlate with different discourse-pragmatic functions and, concomitantly, with different kinds of semantic constraints (e.g. whether or not indefinite DPs are allowed). Examples such as (6) are often considered afterthoughts in the literature, while examples such as (7) are considered antitopics or right dislocations proper (not all authors distinguish between these two functions as they both relate to discourse topics). Examples such as (8) are generally not considered in the literature on right detachments, as they represent the default case of two independent, equally ranked IUs following one after the other.[8] They are included here because they help to make the main point: in the three types of constructions illustrated by (6)–(8), the prosody is a constitutive part of the construction. If you change the prosody, you get a different construction and meaning. In this sense, they are prosody-dependent.

---

[8] As a matter of fact, the sequence of the two IUs in (8) is somewhat remarkable in that both constituents end on a low boundary tone, signalling non-continuation. In narrative monologue, two short clauses such as *She collected the seeds | and her son locked them away* are often phrased in two consecutive IUs of equal rank, but at least the first IU would then often end on a rising boundary tone, signalling continuation.

Note also that it is not possible to fully integrate *das bittere ende* into the preceding IU, as further discussed in reference to example (11) below.

Prosodic groupings thus come in different types with regard to the prosodic packaging. The default case (in monologue at least) is a sequence of equally ranked IUs which, from the prosodic point of view, are basically coordinated and typically present sequences of events, lists of items and the like.[9] The marked case is that of prosodic add-ons that are subordinate to a preceding or following host IU and may differ in the degree of their prosodic independence. Figure 3 illustrates a fairly independent but nevertheless subordinate prosodic add-on; Figure 4 a maximally dependent one.

The preceding discussion was limited to clitic right dislocations. But the two types of prosodic add-ons, i.e. separate but subordinate as in (6) or integrated extension as in (7), are not limited to this morphosyntactic configuration. In German, for example, right-detached constituents may also occur without a directly co-referential constituent in the host clause, as illustrated in (9). Here, the PP *an ihrem fuß* is added in a separate but subordinate IU, as shown in Figure 5. That is, the prosodic packaging is essentially the same as in the case of example (6).

(9) de::r (0.2) dieser schuh passt (0.1)
    whom   this   shoe   fits

    an   ihrem   fuß (0.5)
    on   her   foot
    '(And he wants to marry the woman) whom this shoe fits, on her foot.'

                                                   [Kölnkorpus Aschenp_105]

Figure 5: Waveform and F0 track for the German example in (9).[10]
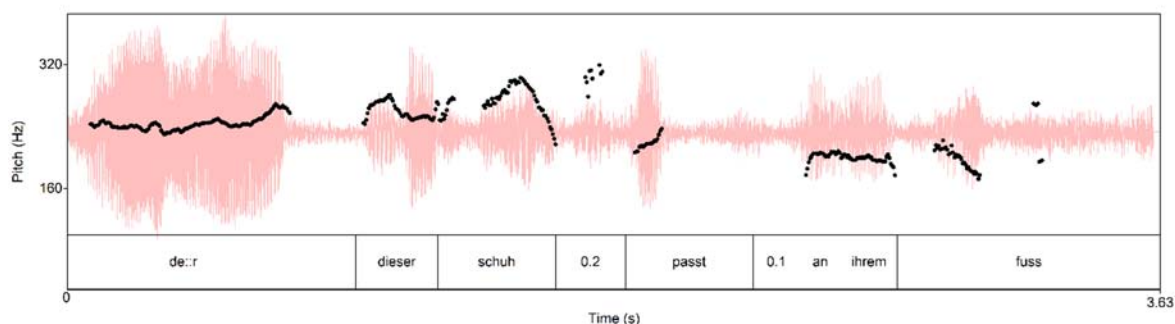
It is a matter of debate how many different prosodic constellations for right-detached constituents may be distinguished (cp., for example, Auer 1996:68-74 for the prosodic packaging options attested for right detachments in German conversations; de Cat 2007:34-62 on French left and right detachments; and Cutfield 2012 on Dalabon right and left detachments involving demonstratives). It is, in fact, very well possible that languages differ with regard to the number and type of packaging options they allow for prosodic add-ons, as will be further illustrated in the following section. However, it is clear that the prosodic options are always much more limited than the morphosyntactic ones. For the latter, different

---

[9] Things are more complex than presented here in that equally ranked IUs typically occur within larger units which, from a phonetic point of view, are linked via a continuously declining pitch base line (cp. Ladd 1988, Schuetze-Coburn 1994, inter alia). See also Crystal's very instructive and detailed discussion of "inter-tone-unit relations" (Crystal 1969:235-252), which also operates with a basic distinction between coordinating and subordinating tone-sequences (Crystal 1969:237 passim).

[10] The F0 traces occurring between *schuh* and *passt* are caused by background noise and can safely be ignored.

types of constructions may be distinguished according to a number of parameters, leading to a great variety of different constructions. Such parameters include argument vs. adjunct function of the detached constituent; presence of a co-referential expression in the host clause; type of co-referential expression (e.g. pronoun, deictic adverbial, DP, PP); and type of detached constituent (e.g. pronoun, adverb, DP, PP, adjectival phrase).

It is also a matter of debate how many different meanings or functions can be distinguished for detached constituents. Typical functions mentioned for right detachments are afterthought (disambiguation), elaboration, antitopic and discourse topic.[11] Importantly, the meanings and functions ascribed to detached constituents essentially depend on the prosodic packaging and not on the morphosyntactic configuration. For example, disambiguations are phrased independently (as in (6)), but highly topical constituents are integrated into the preceding IU (as in (7)). It is for this reason that here, prosodic phrasing is considered to be primary in the definition of prosodic groupings. Prosody is the main carrier of the constructional meaning in these constructions.

3.3 Short summary

With regard to the interface between prosody and syntax, two different constellations need to be distinguished. On the one hand, there are prosodically robust phrase structures which are essentially independent of prosodic phrasing. The boundaries of these phrases are usually aligned with the boundaries of prosodic phrases, but there are systematic exceptions (e.g. heavy constituents or "misaligned" function words), as widely discussed in the literature on clitics and Prosodic Phonology. In the view proposed here, such systematic exceptions demonstrate the prosodic robustness of the phrase structural units that allow for them.

On the other hand, there are prosody-dependent configurations where prosodic phrasing provides the basis for relating syntactic strings consisting of one or more words to each other. One case, hardly ever discussed in the literature, is a sequence of equally ranked, clause-sized IUs, as often occurs in narrative sequences (e.g. [*then a boy comes by*]$_{IU}$ [*riding a bicycle*]$_{IU}$ [*stops below the tree*]$_{IU}$ [*and looks up to the farmer*]$_{IU}$). In another case, much discussed in the literature on detachments, parentheses and appositions, prosodic phrasing marks a syntactic string as being subordinate to a preceding or following one, providing framing or adding information.

The distinction between prosodically robust and prosody-dependent constructions is rarely made explicit in the literature. It is, however, implicit in the fact that most literature on the syntax-prosody interface is exclusively concerned with prosodically robust constructions.[12] The relevance of this distinction is intricately linked to how the interface between syntax and prosody is conceived of. Inasmuch as prosodic phrasing is seen as derivative of syntax, there is no difference at a deeper syntactic level. Both prosodically robust and prosody-dependent constructions are based on syntactic phrase structures and differ only in the way syntactic structure determines prosodic detail (more so in the case of prosody-dependent constructions). Alternatively, syntactic and prosodic structure can be seen to be essentially independent of each other, each based on its own principles.

---

[11] There is little consensus on basic conceptual and terminological issues in this regard, with most terms being used for at least two different kinds of prosodic groupings. See, among others, Lambrecht (1981, 1994), Averintseva-Klisch (2009) and Cutfield (2012) for discussion and examples.

[12] A particularly telling example is Wagner (2015) who extensively discusses prosodic evidence for syntactic structure without mentioning a single prosody-dependent construction.

It should be obvious that the approach advocated in the current paper is clearly on the side of the view that syntactic and prosodic structure are essentially independent of each other. The basic hypothesis is that for prosody-dependent constructions, prosody is primary and morphosyntactic configurations are secondary. It is the prosody that conveys the main meaning of the overall construction (and thus, for example, distinguishes afterthoughts from antitopics). In the case of prosodically robust constructions, on the other hand, the relation is one of (non-)alignment between two essentially independent phrasing levels, with syntactic phrasing providing the main input to meaning composition.

The differences between the two types of prosody-syntax constellations have implications for the architecture of grammar, language processing,[13] diachronic developments and typology. In the next section, we are primarily concerned with implications regarding the latter two fields.

4. On the emergence of prosodically robust phrase structure

The distinction between prosodically robust and prosody-dependent constructions provides a means to clarify the role of prosody in the historical development of (syntactic) phrase structure. Here, we focus on the role of prosodic adjacency as a major prerequisite for the grammaticization of phrase structure. Inasmuch as semantically related units tend to co-occur next to each other, prosodic adjacency is unproblematic in a large number of source constructions for robust phrase structures (section 4.1). However, it appears to be the case that not all prospective co-constituents of an emerging construction may co-occur within the same IU in all languages at all times, as a brief comparison between German and Dyirrbal suggests. To capture such constraints, we introduce *IU-boundedness* as a second basic type of prosody-dependence in section 4.2.

4.1 Prosodic adjacency is a condition for the grammaticization of phrase structure

It is widely assumed that most, if not all, grammaticization processes require that the grammaticizing elements occur adjacent to each other. Thus, for example, Hall (1992:162) mentions "exclusive adjacency" as a precondition for the coalescence of function words with their hosts. That is, in order for function words to become affixes it is necessary that function words and their hosts regularly (and frequently) occur immediately adjacent to each other. This adjacency requirement includes prosodic integration, as noted by Bybee et al. (1990:29).

Adjacency would also seem to be a precondition for the grammaticization of phrase structure. It is a defining characteristic of syntactic phrases that the constituents that make up a phrase normally occur adjacent to each other. This adjacency requirement is only defeasible under specific circumstances (e.g., in the case of heavy DP shift). Thus, for example, if nouns, demonstratives and adjectives are allowed to be separated from each other, as they are in so-called discontinuous DPs, there is no DP as a syntactic constituent in terms of (surface) constituent structure theory.[14] As already mentioned in the previous section, constituent

---

[13] If intonation units are essentially idea units (Chafe 1994), chunking of the message content into IUs must occur very early on in speech planning, immediately following message conception, but preceding all other aspects of grammatical encoding in a model such as the one proposed in Levelt (1989).

[14] This is different for dependency groupings: dependency relations hold regardless of adjacency.

It is an unresolved question whether the grammaticization of phrase structure necessarily presupposes a concurrent grammaticization of a function word which typically serves as a grammatical marker for the emerging phrase (see Himmelmann 1997:155-157 for preliminary discussion). This is clearly so in the case of DPs which require the concurrent grammaticization of determiner-like function words, and similarly for PPs and CPs. However, it is less clearly so in the case of VPs, at least of those which do not have a major slot for

structure tests are means to diagnose the unit behavior of a string of words. Adjacency is a major condition for unit behavior.

With regard to the prerequisites for the grammaticization of phrase structure, adjacency also includes prosodic integration. Given that in the preceding section the property of being independent of prosodic phrasing (i.e. being prosodically robust) was claimed to be a major characteristic of phrase structure proper, this may seem paradoxical at first. However, this is only an apparent contradiction. There is a crucial difference between a prerequisite for a grammaticization process and its outcome. The claim here is that prosodic integration is essential as a prerequisite, but that it is no longer essential once phrase structure proper has emerged. There is a threshold in rigidity that has to be traversed before an emerging phrase structure becomes prosodically robust, and thus recognizable as such without prosodic scaffolding.

Auer (1996) makes a case for the claim that from the point of view of conversational turn taking it makes sense to have different means for relating (groups of) words to each other. Specifically, the availability of both robust phrase structures and prosodic phrasing provides the basis for a common floor-keeping strategy in that it allows speakers to signal continuation that extends beyond the end of a given phrasing type. When a syntactic phrase comes to an end, prosodic phrasing may indicate that the current speaker intends to go on. And vice versa.

There is ample historical evidence for the claim that the grammaticization of phrase structure involves the emergence of an adjacent, and in later stages often fixed, ordering of the co-constituents of the emerging phrase (cp., for example, Himmelmann 1997:140-144, Vincent 1999, van de Velde 2009, Reinöhl 2016, Börjars et al. 2016). As for prosodic integration, direct evidence is difficult to come by given the lack of recordings available for earlier stages in the development of synchronically attested phrase structures. Still, Reinöhl & Casaretto (2018) make a convincing case for the claim that in the absence of prosodic integration, typical grammaticization processes do not take place, thus providing an explanation for why local particles did not grammaticize into adpositions in Indo-Aryan, unlike in most other branches of Indo-European.

Disregarding such cases which, importantly, involve another prosodic phrasing level than the IU, it is likely that prosodic coherence is not really an issue for many standardly recognized phrase structures such as DPs, PPs and CPs. This is due to Behaghel's famous first law (Behaghel 1932:4): "Das oberste Gesetz ist dieses, daß das geistig eng Zusammengehörige auch eng zusammengestellt wird."[15] That is, those words that tend to become co-constituents of a phrasal constituent already tend to occur close to each other. It is therefore likely that they also occur within an IU. Thus, more often than not demonstrative and noun (the source construction for DPs) or a relational noun and its semantic complement (*middle of X, back of X* – one of the source constructions for PPs) will occur within the same IU.

---

auxiliary-type elements. It may also not be the case for some types of the determiner- and adpositionless nominal expressions discussed in section 3.1 of the introduction to this volume inasmuch as these are claimed to show different degrees of phrasal organisation. A strong hypothesis in this regard would be that only phrases with overt function words are the result of grammaticization processes. Other types of syntactic phrases without such function words would not result from grammaticization processes but from other types of syntactic change.

[15] Translation: 'The principal law is this: that what belongs together mentally is also placed close together'. Among many others, see Rijkhoff (2004, 2015) for further details, discussion and references.

A brief look into text collections which indicate prosodic boundaries, such as Heath (1980) for Nunggubuyu (Wubuy) or the appendix to Merlan's (1994) grammar of Wardaman, reveals that this is in fact clearly the case in languages which are widely seen as lacking rigidly structured DPs and other types of rigid phrase structures (see also Croft 2007:23-25 for additional support from Wardaman). To be sure, in these texts one will find many examples in which a demonstrative and a co-referential noun occur in separate IUs. But there is no general restriction against a demonstrative and a co-referential noun occurring in the same IU (cp. example (10) below). In short, the prerequisite for the grammaticization of phrase structure – that grammaticizing co-constituents are allowed to occur within a single IU – is very likely to be met for many configurations from which phrase structures may emerge.

To put this even more strongly, if the preceding argument is correct, the prosodic phrasing prerequisite for the grammaticization of phrase structure proper, i.e. the possibility for prospective co-constituents to occur within the same IU, is fulfilled in most, if not all languages. Hence, differences in prosodic phrasing are an unlikely explanation for the fact that not all languages have evolved robust phrase structures. Other conditions and processes appear to be more relevant in getting the grammaticization process started (cp., for example, Reinöhl 2016 on obligatorification).

4.2 IU-boundedness

The preceding assessment does not mean that no crosslinguistic differences exist with regard to whether a given syntactic unit has to occur within a single IU or may be split across a number of IUs. There are also crosslinguistic differences regarding which syntactic units are allowed to co-occur within a single IU. Such constraints instantiate a type of prosody-dependence different from prosodic groupings. In prosodic groupings, prosodic dependence pertains to the fact that semantic, pragmatic and syntactic relations between (groups of) words depend on the prosodic packaging extending beyond a single IU. In this section, we are dealing with requirements that the constituents of a construction have to, or must not, co-occur within the same IU. For want of a better term, constructions which are subject to such requirements are called *IU-bounded constructions.*

The Australian languages Nunggubuyu and Wardaman just mentioned have in common that all elements that may be part of a nominal expression (noun, demonstrative, quantifier, modifiers of different types, etc.) may form a syntactically complete nominal expression all by themselves. Thus, in example (10) from Nunggubuyu, each of the two words that make up the nominal expression *ngarribiyung ngarrubagi* 'that mother' may be used as a (syntactically) complete nominal expression in an appropriate context.

(10) ngarribiyung[16]   ngarrubagi

     ngarra-ibi-yung   ngarra-uba-gi

     F.SG-mother-3.POSS   F.SG-F.ANAPH-SG

     '(She came along with him) that mother.'   [Heath 1980:46]

Complex nominal expressions in these languages are thus prosody-dependent in that their co-constituents have to occur within a single IU. If *ngarrubagi* were split off from *ngarribiyung* and formed an IU by itself, the overall construction would become a different one (for example, it might be an afterthought or it could be the starting point for a new clause). There

---

[16] Glosses: ANAPH = anaphoric demonstrative; F = feminine; POSS = possessive pronoun; 3 = third person.

is nothing in the grammatical structure of the remaining items (including those preceding *ngarribiyung*) that would indicate that the overall grammatical construction is incomplete. This is different in the case of the German fragmentation example in (3). In this example, *von der* is syntactically incomplete. Hence, a continuation with *strecke* after a prosodic interruption is not framed as an afterthought or as the starting point of a new phrase, but as the continuation of an incomplete construction.

IU-boundedness may also have an exclusionary role. Specifically, the emergence of robust phrase structure appears to correlate with the emergence of constraints on the syntactic structures allowed to co-occur within an IU, at least in European languages. What does not seem to be possible in German is to fully integrate right-detached DPs into the host IU. That is, *die samen* in (7), here repeated as (11), has to be clearly marked as an add-on to an IU that would be complete without it. In particular, the pitch remains flat and intensity is greatly reduced (cp. Figure 4).

(11) ich hab **die   alle** abgemacht die **samen** (0.5)
    I   have them all   removed   the seeds
    'I removed them all the seeds.'   [Kölnkorpus Oma_089]

This is different in Australian languages where we find examples such as (12) from Dyirbal.[17]

(12) **bala-n**   ngamban **gunyu**   **dambun**
    DIST.ABS-II hear.PAST new.ABS   dambun.ABS
    '(It was) that other Dambun (I) heard.'   [Dixon 1972:385, line 19]

In this example, the final nominal expression *gunyu dambun* 'new/other Dambun (k.o. ghost)' is fully integrated into the IU and, unlike in the German examples (6) and (7), is not a prosodic add-on. Compare Figure 4 with Figure 6.

Figure 6: Waveform and F0 track for the Dyirbal example in (12).
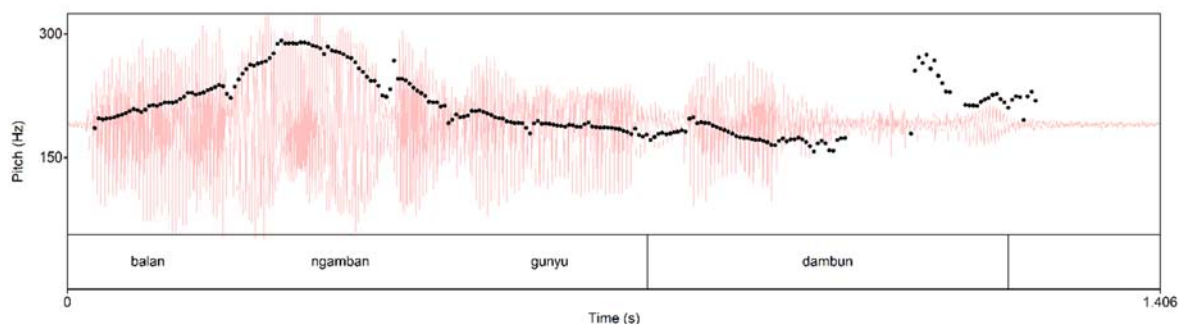
In Figure 6 there is no prosodic break between *ngamban* and *gunyu*. Rather, *gunyu* is part of a continuous fall which reaches its deepest point on *dam* and is followed by a rise on *bun*

---

[17] Glosses: II = noun class II; ABS = absolutive case (unmarked), DIST = distal demonstrative. The recordings of a number of the texts published in the appendix of Dixon (1972) were made available on audio cassette to the author by the Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS). See also Dixon (1972:368).
Nominal expressions in Dyirbal are structurally considerably less flexible than nominal expressions in Nunggubuyu and Wardaman. But the observation that all constituents of a nominal expression may function as its distributional equivalent also applies to Dyirbal.

(which is analyzed as a final rising boundary tone by King 1998). In Figure 4, on the other hand, the final nominal expression *die samen* is arguably also integrated into the preceding IU, but is clearly marked as an add-on. One could remove *die samen* and still end up with a coherent and complete IU.

Another way to state this difference between Dyirbal and German is to say that German lacks discontinuous nominal expressions. In the literature, this observation is often stated without proper attention to prosodic phrasing, which easily leads to confusion. The difference is not that German does not allow the distribution of what one may consider co-constituents of a complex nominal expression across non-adjacent strings. This is perfectly possible as long as the second string is prosodically packaged as an add-on or in an IU of its own. Furthermore, if in the specific case of example (11) this prosodic restriction were to be lifted, the resulting construction would be a clitic doubling construction, which would again differ from the discontinuous nominal expression in the Dyirbal example.

Clitic doubling constructions do not occur in German (**ich hab die alle abgemacht die samen* – where *die samen* is a fully integrated constituent of the IU, i.e. carries a postlexical pitch accent on *sa* and the IU-final boundary tone on *men*). There appears to be a constraint against having two separate, co-referential DPs fully integrated in a single IU in German and all the other European languages that do not allow clitic doubling. This constraint would indeed concern the syntactic category and not the weight or length of the phrases. That is, as assumed in standard constituent structure analyses, *die* and *die samen* in (11) (or *die* and *die schoh* in (6)) are both proper phrasal constituents even though the first one is only represented by a pronoun.

This state of affairs would contrast, on the one hand, with the discontinuous nominal expression from Dyirbal in (12) where different parts of what is functionally and, if one follows the analysis of similar constructions in Jaminjung by Schultze-Berndt and Simard (2012), also structurally a single nominal expression may co-occur fully integrated into a single IU. On the other hand, it would contrast with a clitic doubling construction such as Rio Platense Spanish *Lo vimos a Juan* (him we-saw *a* Juan) 'we saw Juan', where, if our speculation is correct, the clitic pronominal does not instantiate a phrasal constituent but a function word. It seems possible that clitic doubling constructions historically derive from right-detached constructions where the final DP is realized prosodically as an integrated extension as in (11), but this is a topic for another investigation (see Haig 2018 for some pertinent observations).

The preceding considerations are fully compatible with the typology of structuring options for nominal expressions proposed in the introduction to this special issue. If prosodic integration within a single IU is the precondition for the grammaticization of phrase structure, we may expect various degrees of rigidity in the structure of nominal expressions that may occur in a single IU. Once a certain threshold in rigidity has been traversed, an emerging phrase structure becomes prosodically robust, and thus recognizable as such without prosodic scaffolding.

5. Conclusion

With regard to the relation between syntactic and prosodic phrases, the present article has argued for a distinction between prosodically robust phrase structures and prosody-dependent constructions. Prosodically robust phrase structures are the result of a grammaticization

process. For this grammaticization process, prosodic phrasing only plays an ancillary role in that the grammaticization of robust phrase structure presupposes that the co-constituents of the emerging phrasal constituent regularly occur within one IU. Once phrase structure proper has been grammaticized and hence has properties of a formal gestalt that exists independently of prosodic scaffolding, prosodic and syntactic phrasing provide alternative ways of signaling the relation between (strings of) words. While the boundaries of prosodic and syntactic phrases are often aligned with each other, precise alignment is not necessary and there is considerable flexibility with regard to the size of the syntactic units presented within a single IU.

Prosody-dependent constructions, on the other hand, essentially depend on the prosodic cues that delimit their boundaries. There are two basic types of such constructions: prosodic groupings and IU-bounded constructions. The latter require that the co-constituents of a construction co-occur within the same IU. Prosodic groupings extend beyond the bounds of a single IU. They may involve a combination of two or more equally ranked IUs, or they involve the combination of one superordinate with one or more subordinate prosodic phrases. Subordinate phrases are prosodic add-ons to the host IU. There are different subtypes of prosodic add-ons relating primarily to the degree of their prosodic independence. In section 3.2, for example, a distinction is made between an add-on that appears in an IU of its own and one which is integrated into the host IU (called IU extensions in section 3.2). It is very well possible that languages distinguish a different number of subtypes of prosodic add-ons, but their number will generally be fairly small.

There are a number of implications that follow from the distinction between prosodically robust phrase structures and prosody-dependent constructions. Only one of these, pertaining to the role of prosody in the emergence of phrase structures, is explored in section 4. An implication not further explored here pertains to the usefulness of prosodic diagnostics for syntactic constructions. A prototypical example are serial verb constructions, specifically core layer or co-ranked serialization. Prosodic integration is regularly mentioned as a defining feature of these constructions, distinguishing them, inter alia, from sequences of clauses. From the point of view argued here, two issues are of interest. First, the ability to occur within a single IU can only be a necessary but never a sufficient criterion for a phrasal construction (cp. Unterladstetter 2019). IUs in all languages may contain multiple clauses, VPs or verbs (cp. Ross et al. 2016). Second, as phrase structures proper, serial verb constructions would have to be prosodically robust, i.e. recognizable as such across two (or more) consecutive IUs. The serial verb literature does not properly address this issue, but it seems likely that this condition is rarely if ever met in the case of (co-ranked) serial verb constructions.

clarify important aspects of the ideas presented here. Many thanks also to Katherine Walker for improving English grammar and style, and to Katharina Gayler and Sonja Riesberg for crosschecking the final version.

References

Altmann, Hans. 1981. *Formen der »Herausstellung« im Deutschen*. Tübingen: Niemeyer.

Auer, Peter. 1991. Vom Ende deutscher Sätze. *Zeitschrift für Germanistische Linguistik* 19:139-157.

Auer, Peter. 1996. On the prosody and syntax of turntaking. In: Elizabeth Couper-Kuhlen & Margret Selting (eds), *Prosody in conversation: Interactional studies*. Cambridge: CUP, 57-100.

Averintseva-Klisch, Maria. 2009. *Rechte Satzperipherie im Diskurs. NP-Rechtsversetzung im Deutschen*. Tübingen: Stauffenburg.

Behaghel, Otto. 1932. *Deutsches Syntax 4: Wortstellung*. Heidelberg: Winter.

Bhatt, Christa & Markus Lindlar (eds). 1998. *Alles Kölsch*. Eine Dokumentation der aktuellen Stadtsprache in Köln. Bonn: Bouvier Verlag.

Börjars, Kersti, Pauline Harries & Nigel Vincent. 2016. Growing syntax*: the* development *of a* DP *in* North Germanic. *Language* 92, e1–37.

Bruening, Benjamin. 2020. The head of the nominal is N, not D: N-to-D Movement, Hybrid Agreement, and conventionalized expressions. *Glossa: a journal of general linguistics*, 5(1), p.15. DOI: http://doi.org/10.5334/gjgl.1031.

Bybee, Joan L., W. Pagliuca, and R.D. Perkins. 1990. On the asymmetries in the affixation of grammatical material. In: William Croft et al. (eds), *Studies in typology and diachrony: Papers presented to Joseph H. Greenberg on his 75th birthday.*. Amsterdam: Benjamins, 1-42.

Chafe, Wallace L. 1994. *Discourse, Consciousness, and Time*. Chicago: The University of Chicago Press.

Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.

Couper-Kuhlen, Elizabeth. 1986. *An Introduction to English prosody*. Tübingen: Niemeyer.

Croft, William. 1995. Intonation units and grammatical structure. *Linguistics* 33:839-882.

Croft, William. 2007. Intonation Units and Grammatical Structure in Wardaman and in Cross-linguistic Perspective. *Australian Journal of Linguistics* 27: 1-39. DOI: 10.1080/07268600601172934.

Crystal, David. 1969. *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press.

Cutfield, Sarah. 2012. *Demonstratives in Dalabon: a language of south-western Arnhem Land*. Ph.D. Thesis, Monash University, Melbourne.

de Cat, Cécile. 2007. *French Dislocation. Interpretation, Syntax, Acquisition*. Oxford: OUP.

Dehé, Nicole. 2014. *Parentheticals in Spoken English: The Syntax-Prosody Relation*. Cambridge: Cambridge University Press.

Dixon, R.M.W. 1972. *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press.

DuBois, John W., Stephan Schuetze-Coburn, Danae Paolino & Susanna Cumming. 1992. *Discourse transcription*. Santa Barbara: UCSB, Department of Linguistics (Santa Barbara Papers in Linguistics, 4).

Elfner, Emily. 2018. The syntax-prosody interface: current theoretical approaches and outstanding questions. *Linguistics Vanguard* 1, 20160081. doi: https://doi.org/10.1515/lingvan-2016-0081

Hall, Christopher J. 1992. *Morphology and mind: A unified approach to explanation in linguistics*. London: Routledge.

Haig, Geoffrey. 2018. The grammaticalization of object pronouns: Why differential object indexing is an attractor state. *Linguistics* 56:781-818.

Heath, Jeffrey. 1980. *Nunggubuyu myths and ethnographic texts*. Canberra: AIAS.

Himmelmann, Nikolaus P. 1997. *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Tübingen: Niemeyer

Himmelmann, Nikolaus P. 2014. Asymmetries in the prosodic phrasing of function words: Another look at the suffixing preference. *Language* 90:927-960.

Himmelmann, Nikolaus P. 2015. Notes on 'noun phrase structure' in Tagalog. In Jens Fleischhauer, Anja Latrouite & Rainer Osswald (eds), *Explorations of the Syntax-Semantics Interface*. Düsseldorf: dup, 315-337.

Himmelmann, Nikolaus P., Meytal Sandler, Jan Strunk & Volker Unterladstetter. 2018. On the universality of intonational phrases in spontaneous speech – a cross-linguistic interrater study. *Phonology* 35:207-245.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, Mass.: The MIT Press.

King, Heather B. 1998. *The declarative intonation of Dyirbal: an acoustic analysis*. Munich: Lincom Europa.

Ladd, D. Robert. 1988. Declination 'reset' and the hierarchical organization of utterances. *Journal of the Acoustical Society of America* 84:530-544.

Ladd, D. Robert. 2008. *Intonational phonology*. 2nd ed. Cambridge: Cambridge University Press.

Lambrecht, Knud. 1981. *Topic, Antitopic and Verb Agreement in Non-Standard French*. Amsterdam: John Benjamins.

Lambrecht, Knud. 1994. *Information structure and sentence form : topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

Levelt, Willem J.M. 1989. *Speaking: From intention to articulation*. Cambridge, Mass.: The MIT Press.

McCawley, James D. 1989. INFL', Spec, and other fabulous beasts. *Behavioral and Brain Sciences* 12:350-352.

Matsumoto, Kazuko. 2000. Japanese intonation units and syntactic structure. *Studies in Language* 24:515–564.

Matthews, Peter H. 2007. *Syntactic Relations. A Critical Survey*. Cambridge: Cambridge University Press.

Merlan, Francesca C. 1994. *A Grammar of Wardaman. A language of the Northern Territory of Australia*. Berlin & New York: Mouton de Gruyter.

Michaelis, Laura & Knud Lambrecht. 1996. Toward a construction-based theory of language function: the case of nominal extraposition. *Language* 72: 215-247.

Nespor, Marina & Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris.

Park, Joseph Sung-Yul. 2002. Cognitive and interactional motivations for the intonation unit. *Studies in Language* 26:637–680.

Pawley, Andrew & Frances Syder. 2000. The one-clause-at-a-time hypothesis. In: Heidi Riggenbach (ed.), *Perspectives on Fluency*. Ann Arbor: U. Michigan Press, 163-198.

Pullum, Geoffrey. 1985. Assuming some Version of X-bar Theory. *Chicago Linguistic Society* 21:323-353.

Reinöhl, Uta. 2016. *Grammaticalization and the Rise of Configurationality in Indo-Aryan*. Oxford: Oxford University Press.

Reinöhl, Uta & Antje Casaretto. 2018. When grammaticalization does NOT occur. Prosody-syntax mismatches in Indo-Aryan. *Diachronica* 35(2). 238-276.

Richards, Norvin. 2016. *Contiguity Theory*. Cambridge, MA: MIT Press.

Rijkhoff, Jan. 2004. *The noun phrase*. Oxford: Oxford University Press.

Rijkhoff, Jan. 2015. *Word order*. In: James D. Wright (ed.), *International encyclopedia of the social & behavioral sciences*. 2nd edition, Vol. 25, 644–56. Oxford: Elsevier.

Ross, Bella, Janet Fletcher & Rachel Nordlinger. 2016. The alignment of prosody and clausal structure

in Dalabon. *Australian Journal of Linguistics* 36:52-78.

Scheer, Tobias. 2010. *A Guide to Morphosyntax-Phonology Interface Theories: How Extra-Phonological Information is Treated in Phonology since Trubetzkoy's Grenzsignale*. Berlin: de Gruyter.

Schuetze-Coburn, Stephan. 1994. *Prosody, syntax, and discourse pragmatics: assessing information flow in German conversation*. Ph.D. dissertation, University of California, Los Angeles.

Schultze-Berndt, Eva. 2002. Constructions in language description. *Functions of Language* 9:267-308.

Schultze-Berndt, Eva & Candide Simard. 2012. Constraints on noun phrase discontinuity in an Australian language: The role of prosody and information structure. *Linguistics* 50: 1015–1058.

Selkirk, Elisabeth. 1984. *Phonology and Syntax: The relation between sound and structure*. Cambridge, Mass.: The MIT Press

Selkirk, Elisabeth. 2011. The syntax– phonology interface. In Goldsmith, John A., J. Riggle & A. Yu (eds.), *The Handbook of Phonological Theory*. 2nd edition. Oxford: Blackwell, 435 – 484.

Selting, Margret. 1995. *Prosodie im Gespräch. Aspekte einer interaktionalen Phonologie der Konversation*. Tübingen: Niemeyer.

Tao, Hongyin. 1996. *Units in Mandarin Conversation. Prosody, discourse, and grammar*. Amsterdam: Benjamins.

Truckenbrodt, Hubert. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30:219-255.

Unterladstetter, Volker. 2019. *Multi-verb constructions in Eastern Indonesia*. Berlin: Language Science Press.

van de Velde, Freek. 2009. *De nominale constituent. Structuur en geschiedenis*. Leuven: Universitiere Pers Leuven.

Vincent, Nigel. 1999. The evolution of c-structure: prepositions and PPs from Indo-European to Romance. *Linguistics* 37:1111-1153.

Wagner, Michael. 2010. Prosody and recursion in coordinate structures and beyond. *Natural Language & Linguistic Theory* 28:183-237.

Wagner, Michael. 2015. Phonological Evidence in Syntax. In Tibor Kiss & Artemis Alexiadou (eds), *Syntax–Theory and Analysis. An International Handbook*. Berlin: de Gruyter Mouton, 1154–1198.