Syntax, Information Structure, Embedded Prosodic Phrasing and the Relational Scaling of Pitch Accents

Submitted to Erteschik-Shir, Nomi and Lisa Rochman (eds) The Sound Patterns of Syntax. Oxford University Press.

Caroline Féry (caroline.fery@googlemail.com)

Potsdam, July 2008, 1st draft.

1. Background

1.1 Prosodic domains and syntax

A breakthrough in the long history of the theoretical approaches to how syntax shapes prosody was achieved by the emergence of prosodic hierarchies, as proposed by Selkirk (1984), Nespor & Vogel (1986) and many others after them. Prosodic hierarchies capture the insight that morpho-syntactic units are mapped to prosodic units of different sizes, even if the mapping is not strictly isomorphic. A grammatical word, for instance, often forms a Prosodic Word, and some morphological operations, like reduplication or hypochoristic formation in many languages, can only be fully understood if their prosodic structure is taken into account (McCarthy & Prince 1990). At the top of the prosodic hierarchy, sentences correspond to Intonation Phrases (Liberman 1978, Pierrehumbert 1980, Liberman & Pierrehumbert 1984) and are assigned intonational patterns.

Phonologists largely agree on units like Prosodic Words and Intonation Phrases, but the intermediate prosodic domains have been a matter of debate. Most researchers assume two levels of prosodic phrasing between Prosodic Word and Intonation Phrase, and these have been given a variety of names, as for instance 'Minor Phrase' and 'Major Phrase' (Poser 1984, Selkirk 1986), 'Accent Phrase' and 'Intermediate Phrase' (Beckman & Pierrehumbert 1986, Gussenhoven 2004, Jun 2005), or 'Clitic Group' and 'Phonological Phrase' (Nespor & Vogel 1986). Together with a restrictive view of what is allowed in the prosodic mapping from syntax to prosodic structure, like the Strict Layer Hypothesis (Selkirk 1984, Nespor & Vogel 1986), which forbids recursive structure, the assumption of a maximum of two layers of phrasing can be interpreted as a prohibition on long sentences.

But of course, in the same way as syntax cannot restrict sentences to a certain length—there is no way of forcing a sentence to have, say, maximally five embedded

clauses—there should also be no way of restricting the number of prosodic domains that a sentence may have. For this reason, the prosodic phrasing resulting from the syntax-prosody mapping should be recursive and unconstrained. This mapping results in prosodic domains that are called p-phrases (for prosodic phrases), and that can be embedded into each other. A few authors have proposed getting rid of the two-level-problem by eliminating the restrictive prosodic hierarchy, at least for those levels which are situated at the interface between syntax/semantics and prosody. Wagner (2005) proposes calling all levels of prosodic structure 'feet,' and Ito & Mester (2007) call the levels higher than the metrical foot 'phonological phrase' or just 'phrase.' Ito & Mester distinguish between 'minimal phrase,' 'phrases' and 'maximal phrase.' These three levels of phrasing may present different properties. Only the non-minimal and non-maximal phrases are recursive.

The present paper assumes recursive phrasing at all levels of the prosodic hierarchy, starting at the Prosodic Word. In particular, it assumes a recursive phrasing pattern of p-phrases and intonation phrases (called here i-phrases), but it does not propose to eliminate the prosodic hierarchy.

We will not be concerned too much about the details of how to construct prosodic phrasing from the syntactic structure. To name just a few successful theories which have elaborated formulations of this mapping, consider 'relation-based' theories (Nespor & Vogel 1986), 'edge-based' theories (Selkirk 1986), alignment in Optimality Theory (Truckenbrodt 1999, Selkirk 2000, Féry & Samek-Lodovici 2006), and minimalist phase and spell-out (Ishihara 2007, Kahnemuyipour 2004, Kratzer & Selkirk 2007). The main idea behind all accounts is that syntactic categories are isomorphic with prosodic phrases, or expressed differently, prosodic domains correspond to syntactic phrases (relation-based and phase) or syntactic edges (edge-based and alignment). For English, or German, the language that will occupy us below, prosodic phrasing correlates with pitch accents. Every p-phrase is assumed to have a head, which is instantiated as a pitch accent, and every pitch accent is the head of a p-phrase.

The same few problems are recurrent in all theoretical approaches to prosodic phrasing. These problems are related to the fact that prosodic phrases are projected from the surface syntactic structure, but that this syntactic level does not contain all the necessary information to explain the observed accent pattern.

First, consider German sentences with unergative intransitive verbs. In this kind of sentence, both the subject and the verb are accented, and thus in separate p-phrases, as

shown in (1a). Sentences with unaccusative intransitive verbs, in contrast, have only one accent on the subject and none on the verb. This is illustrated in (1b).

(1) a. [Ein JUNGE]$_P$ [TANZT]$_P$
a boy dances
'A boy dances.'
b. [Die DIVA ist gestorben]$_P$
The diva is died
'The diva has died.'

Since these sentences have the same surface syntactic structure, it has been assumed that deep syntactic properties are relevant for sentence accent assignment (Krifka 1984, Kratzer 1988, Diesing 1988). The subject is VP-internal in (1b), but VP-external in (1a).[1] In a minimalist model, in which each phase is a spell-out domain with its own pitch accent, it must be assumed that the verb plus subject are a single phase in the case of unaccusatives, but are spelled out in two phases in sentences with unergative verbs.

A second thorny aspect of the phrasing and subsequent accent pattern has to do with the distinction between argument and adjunct (see Gussenhoven 1992). Especially in locative prepositional phrases, but also in other types of prepositional or adverbial phrases, it is sometimes difficult to assess the argumental or adjunctive nature of phrases. Consider the German examples in (2). In (2a), the locative is an argument, but in (2b), it is an adjunct (see also Krifka 1984 for such pairs). The distinction is essential because a verb is part of the p-phrase of an adjacent argument, and is consequently not accented, whereas the same is not true in case of an adjunct. In this latter configuration, both the adjunct and the verb are maximal projections and are phrased separately.

(2) a. [MORITZ]$_P$ [hat in STUTTGART übernachtet]$_P$
Moritz has in Stuttgart spent-the-night
'Moritz spent the night in Stuttgart.'
b. [MORITZ]$_P$ [has in STUTTGART]$_P$ [GESUNGEN]$_P$
Moritz has in Stuttgart sung
'Moritz sang in Stuttgart.'

[1] Or the subject is a specifier in the VP, whereas the object is a complement in the VP.

The third difficult case comes from aspects of information structure which change the pitch accent structure of a sentence. Some authors treat the effects of syntax and information structure on prosody in the same way in that information structure primarily changes the phrasing of sentences. The change in accent structure is then an indirect effect of the changed p-phrasing (Gussenhoven 1992, Truckenbrodt 1999). In a minimalist approach, this view of phrasing means that prosodic phrasing is not mapped only to phases, but that the information structure also projects phases and spell-out domains. The changed p-phrasing because of information structure is illustrated in (3) with English examples. In (3a) the sentence is all-new and has two accents, one on the subject and one on the object. In (3b), only the subject is focused. As a result, the phrasing has changed because there is a unique accent on the subject, and none on the object. All accounts of prosodic phrasing which assume that every p-phrase is obligatorily headed by a pitch accent have to change prosodic phrasing when accent structure is modified in this way.

(3) a. {What happened?}
[MAX]$_P$ [stole a CHICKEN]$_P$
b. {Who stole a chicken?}
[MAX$_F$ stole a chicken]$_P$

In this paper, I propose separating the effects of syntax from those of information structure. Only syntax influences phrasing, and information structure determines the presence and the height of pitch accents. In some cases, pitch accents are just not realized, and a p-phrase can exist without a pitch accent. As a result, example (3) always has the phrasing shown in (3a), regardless of the information structure and pitch accents.

Some problems related to the phrasing of sentences with different information structure patterns remain that have to do with the contextual framework of the sentences. For example, the same sentence can be thetic or categorical, depending on the context in which it is uttered. If the subject of a presentational sentence is a topic, like in (4a), both the subject and the verb are in separate p-phrases, and both have an accent. But if the whole sentence expresses a unique event, as in (4b), only the subject has an accent, because the subject and the verb are part of the same p-phrase. Since these sentences are both all-new, the difference in accent structure is truly due to a difference in phrasing. In other words, the deaccenting of the particle *durch* is not due to givenness, as in example

(3), in which the chicken had already been mentioned in the context, and was deaccented for this reason. Similarly to (1), in which unergative and unaccusative verbs have different syntactic structures, thetic sentences must be syntactically distinct from categorical sentences. But I know of no syntactic approach which distinguishes between thetic and categorical sentences.

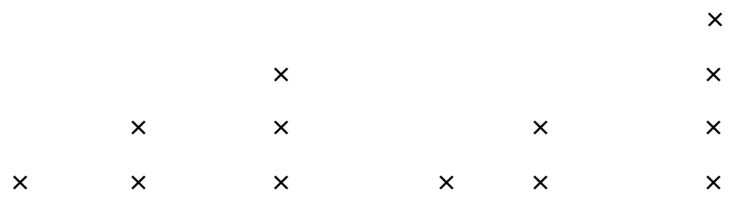
(4) a. [Ein ZUG]$_P$ [fährt DURCH]$_P$ (und ein Auto muss an der Ampel warten.)
A train is-passing through (and a car must wait at the traffic light.)
b. (Achtung auf Gleis 1.) [Ein ZUG fährt durch]$_P$
Attention on platform 1. A train is passing through

The theoretical status of pitch accents as heads of p-phrases has been instantiated in the form of a metrical structure which calculates the difference in strength of the metrical positions from their level of embedding in a tree or in a grid (Liberman 1978, Liberman & Prince 1977, Selkirk 1984, Halle & Vergnaud 1987). In some of these approaches (see for instance Selkirk 2002), levels of the metrical structure may strictly correspond to prosodic domains, as shown in (5). The head of an Intonation Phrase (IP) has a stronger metrical position than the head of a p-phrase, which is itself higher than the head of a Prosodic Word (PW). If the number of levels in the prosodic structure is invariable, a one-to-one correspondence between metrical beats and prosodic domains can be established.

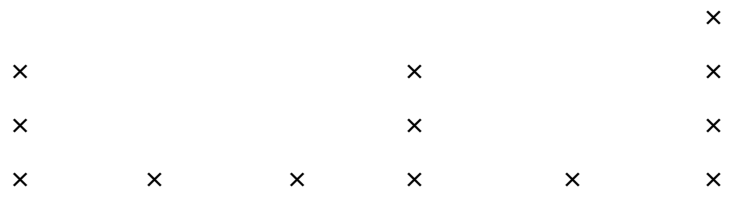
(5) (×) IP
(×) (×) p-phrase
(×) (×) (×)(×) accent domain
(×)(×) (×)(×) (×)(×)(×) (×) PW
Princess Diana's sudden death has been the source of many speculations

However, if p-phrasing is recursive, as assumed here, such a correspondence is not necessary. The height of metrical beats is related to the number of embeddings in the p-phrasing. This is illustrated in (6) with names grouped in different ways (see Wagner 2005 for examples of this kind). In such a case, it is not possible to attribute a specific level of the prosodic hierarchy to a specific grouping. Doing so would inflate the number of prosodic domains in an uncontrollable way.

(6)



a. (((Lena and Arno)$_P$ and Bill)$_P$ ((and Tom and Anny)$_P$ (and Sam)$_P$)$_P$)$_P$



b. (((Lena)$_P$ (and Arno and Bill and Tom)$_P$)$_P$ (and Anny and Sam)$_P$)$_P$

To conclude this section, prosodic domains are mapped from syntactic phrases, but this mapping does not necessarily correlate with specific levels of prosody. Recursivity of p-phrases is assumed, which allows a finer-grained scaling of pitch accents, as shown in the next section.

1.2 The height (and strength) of pitch accents

A problem which has only seldom been addressed in the relevant literature is how to calculate the fundamental frequency (f0) value of accents based on their prosodic and metrical positions (but see Pierrehumbert 1980, Liberman & Pierrehumbert 1984, Truckenbrodt 2004 for proposals involving simple structures). Consider nuclear stress. Since Chomsky & Halle (1968), phonologists and syntacticians regularly mention that the last accent in the sentence is the strongest one of the sentence, and that it is the nuclear stress.[2] This is certainly true for an accent standing for a narrow contrastive focus, especially if it is an early constituent in the sentence (see (3b)). In this case, the pitch accent is the last one, and postfocal material is flat and low, which gives an impression of extra prominence on the accent. But things are different when the sentence is 'all-new,' that is when no constituent in the sentence is particularly emphasized, as in (3a) or (5). In this case, the nuclear stress is generally the pitch accent with the lowest frequency, the smaller pitch range and the weakest acoustic energy. The reason for this is to be found in the downstep pattern of pitch accents, which reduces each pitch accent relative to the

[2] Gussenhoven (1992) and Selkirk (2000) deny the presence of nuclear stress in all-new sentences.

preceding one (see Liberman & Pierrehumbert 1984 for English and Truckenbrodt 2004 for German).

To account for this effect, I propose that p-phrases have an abstract range inside of which accents are scaled (see Bruce 1977, Clements 1981, Ladd 1990 for similar proposals for different languages). In the unmarked case, a sequence of p-phrases at the same level of prosodic phrasing has pitch ranges organized in a downstepped pattern, as illustrated in Figure 1. Since the range of a p-phrase is narrower than the range of a preceding p-phrase in the same i-phrase, the pitch accent heading it is lower than the pitch accent preceding it. The reduced prominence of the last accent (the nuclear stress) in an all-new sentence is just a consequence of this pattern. Pitch accents are indicated with the help of a convex form.

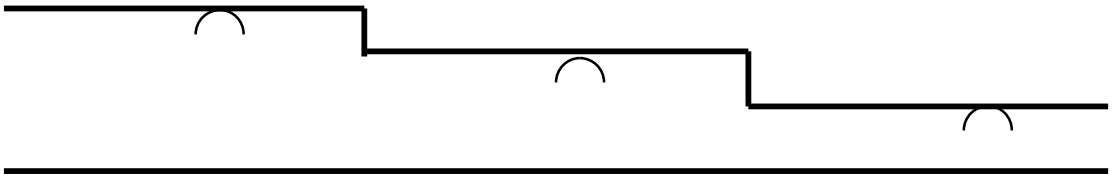


Fig.1 Downstep pattern of unembedded p-phrases

Since the p-phrasing is recursive, every p-phrase can itself contain p-phrases, which are also in a downstep relationship to each other. The head pitch accents are thus scaled inside of these embedded p-phrases. A similar proposal of embedding downstepped regions has been made by a number of researchers (Ladd 1990, van den Berg, Gussenhoven & Rietveld 1992, Truckenbrodt 2007, Féry & Truckenbrodt 2005). This is illustrated in Figure 2.

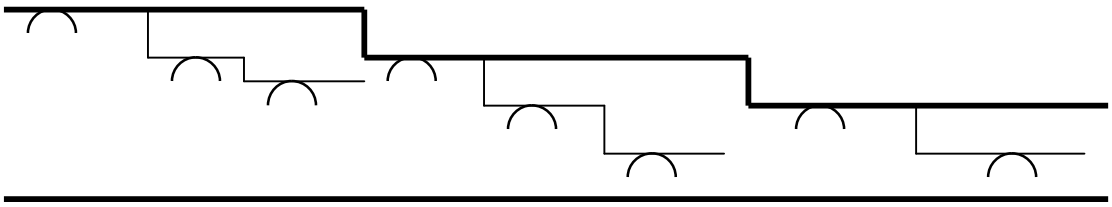


Fig.2 Downstep pattern of embedded p-phrases

Information structure, like focus or givenness, enlarges or reduces the range of a prosodic phrase. A narrow focus has the effect of raising the top line of the corresponding p-phrase, and a given constituent has the effect of lowering it. This is illustrated in Figure 3 and Figure 4, respectively.

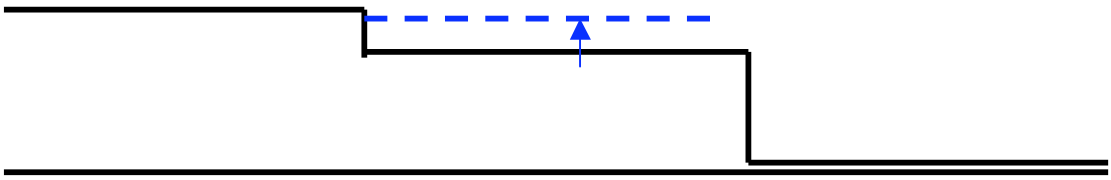


Fig.3 Raising of the top line of a p-phrase because of narrow focus

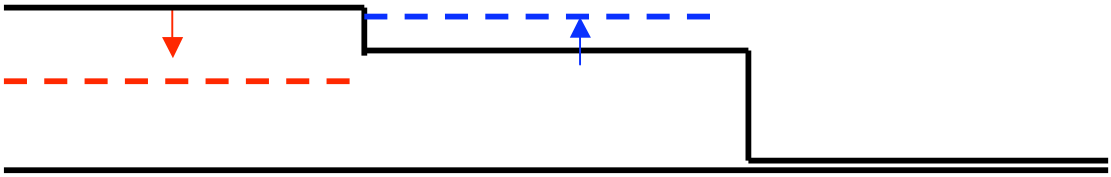


Fig.4 Lowering of the top line of a p-phrase because of givenness

As far as givenness is concerned, a difference is made between pre- and postnuclearity. If a given p-phrase appears before the nuclear accent, its range is narrowed, but a pitch accent can still be realized (see Fig. 4). In the postnuclear region, however, the range is completely compressed, and no pitch accent can be realized anymore.

This model of pitch accent scaling makes a number of predictions:

1) Downstep: An early pitch accent is higher than a later one in the same sentence, everything else being equal. Embedding of pitch ranges into each other accounts for finer differences.

2) Reset: A later accent can be higher than a preceding one if the preceding accent is the head of a more deeply embedded p-phrase.

3) Relative height: The scaling of pitch accents is relative. This means that a pitch accent height may only be raised or lowered as compared to pitch accents in the same intonational domain.

In the next section, some experimental data are discussed that confirm these predictions.

2. Experimental results

2.1 Downstep occurs in all-new sentences

The first prediction claims that pitch accents at the same level of prosodic phrasing are in a downstep relationship to each other, as illustrated in Figs. 1 and 2. This has been shown a number of times for a sequence of arguments in simple syntactic structures or in lists, for

English, German and other languages (see, for example, Liberman & Pierrehumbert 1984 and Ladd 1990 for English, van den Berg, Gussenhoven & Rietveld 1992 for Dutch, Truckenbrodt 2004 for Southern German).

Féry & Kügler (2008) show that in a German simple syntactic sentence, downstep is just one option of how to realize several accents in a sequence. Another one is that the last accent is upstepped and is thus much higher than it would be if downstep had happened regularly. We explain this result with an optional rule of H-raising (see, for instance, Laniran & Clements 2003 for H-raising in Yoruba, and Xu 1999 in Chinese). Notwithstanding the occurrence of H-raising in part of the data, downstep is considered the default realization of a Standard German all-new sentence.

2.2 Reset at a p-phrase boundary

The second prediction posits that the second of a sequence of two accents can be higher than the first one if they belong to different prosodic domains. This effect has been called 'reset' by Liberman & Pierrehumbert (1984). The constellation has been illustrated in Fig. 2, and is also visible in Figure 5 from van den Berg, Gussenhoven & Rietveld (1992). This figure shows that when a larger utterance is divided into two shorter i-phrases, each of them contains downstepped accents, and the first high tone of the second i-phrase is higher than the last tone of the first i-phrase.

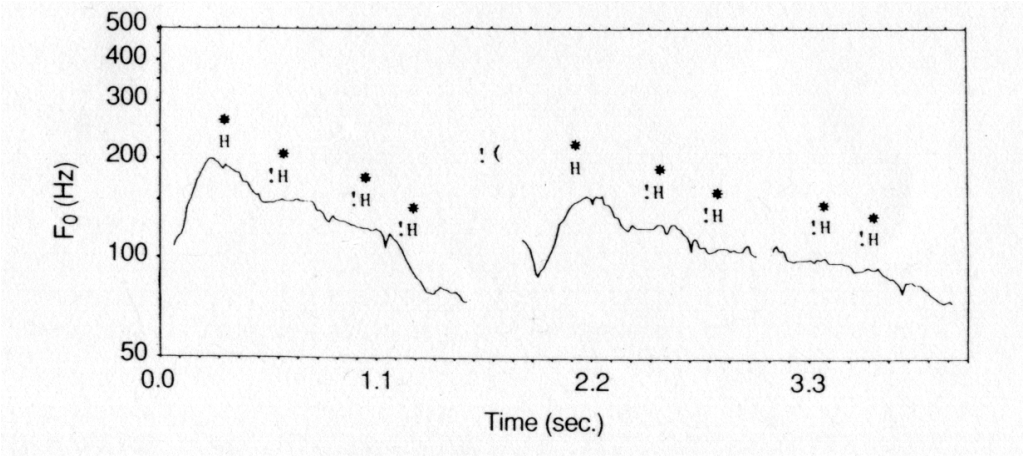


Fig.5 Partial reset and embedded downstep in the Dutch utterance (*Merel, Nora, Leo, Remy), en (Nelie, Mary, Leendert, Mona en Lorna*). From van den Berg, Gussenhoven & Rietveld (1992)

Féry & Truckenbrodt (2005) reproduced for German an experiment by Ladd (1990), who showed that a sequence of three syntactically and semantically related English sentences are in a downstep and/or reset relationship, depending on how their internal syntactic and prosodic structure looks. In Féry & Truckenbrodt, two conditions were examined in a production experiment with the patterns in (7) and (8).

(7) First condition: A while [B and C]

{Why does Anna think that craftsmen have more expensive cars than musicians?}

[Weil der Maler einen Jaguar hat]$_A$, [[während die Sängerin einen Lada besitzt]$_B$, und [der Geiger einen Wartburg fährt]$_C$]

'Because the painter has a Jaguar, while the singer owns a Lada, and the violinist drives a Wartburg.'

(8) Second condition: [A and B] while C

b. {Why does Anna think that musicians have less expensive cars than craftsmen?}

[[Weil die SÄNGERIN einen LADA besitzt]$_A$, [und der GEIGER einen WARTBURG fährt]$_B$], [während der MALER einen JAGUAR hat]$_C$

'Because the singer owns a Lada, and the violinist drives a Wartburg, while the painter has a Jaguar.'

The difference between the prosodic structures of the two conditions is illustrated in Figure 6. In the first condition, B and C form a constituent together, and in the second condition, it is A and B which are grouped into a single constituent. In both conditions, the three sentences form a prosodic constituent together, so that the sentence standing alone is also in a relevant scaling relationship to the other two. I assume a recursive structure: all sentences are i-phrases, the grouping of two sentences is also an i-phrase, and the whole utterance is again an i-phrase.[3]

[3] This assumption differs from the pattern presented in Féry & Truckenbrodt, in which we were more traditional in avoiding recursivity of intonation phrases.

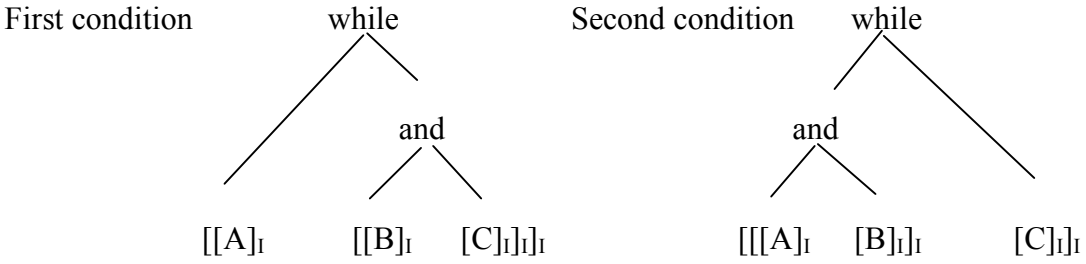


Fig.6 Two conditions in the experiment reported in Féry & Truckenbrodt (2005)

The tonal structure of a sentence of the first condition is shown in (9). Important for the pitch scaling is the value of the first H tone in each sentence.

(9) {Why does Anna think that sportsmen have less expensive cars than craftsmen?}

L*H L*H H_I

[[Weil [der RINGER]$_P$ [[einen LADA]$_P$ besitzt]$_P$]$_I$

L*H L* H H_I

[[während [der MALER]$_P$ [[einen JAGUAR]$_P$ fährt]$_P$]$_I$

L*H H* L L_I

[und [der WEBER]$_P$ [[einen DAIMLER]$_P$ hat]$_P$]$_I$]$_I$]$_I$

'Because the wrestler owns a Lada, while the painter drives a Jaguar and the weaver has a Daimler.'

The production experiment was conducted at the University of Potsdam with five students, native speakers of Standard German, who uttered 32 experimental sentences each. The pattern which emerged from the experiment is that the first condition shows a downstep pattern throughout, as in Figure 7, but the second condition elicited a reset on the C sentence, as shown in Figure 8. The first high tone of this sentence was slightly higher than the first high tone of sentence B. Moreover, this tone was much higher than it was in the first condition.

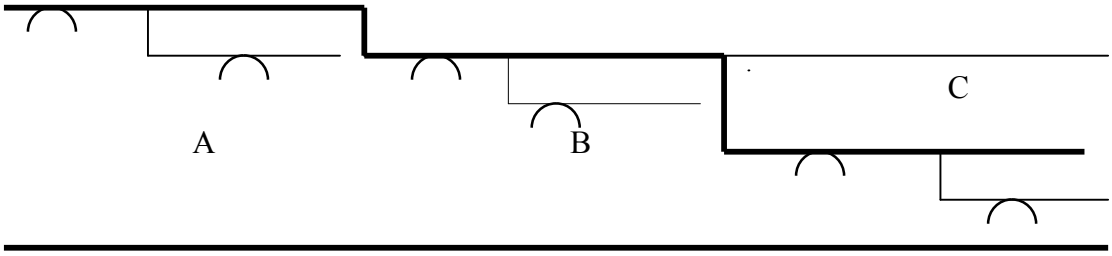


Fig.7 Result of the production experiment of Féry & Truckenbrodt (2005) for the first condition

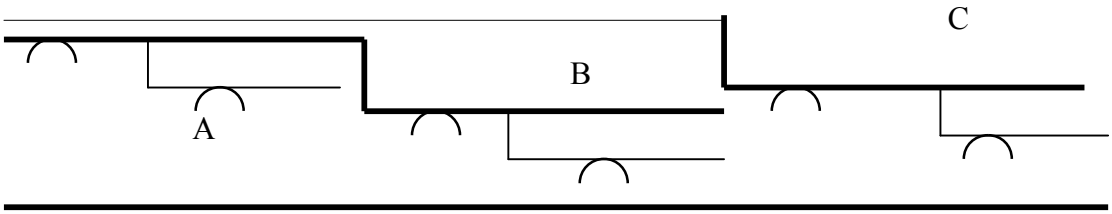


Fig.8 Result of the production experiment of Féry & Truckenbrodt (2005) for the second condition

In short, downstep and reset both play a role in German, and in order to calculate the f0 value of pitch accents in all-new sentences, it is necessary to take both into consideration. A model like the one illustrated in Figs. 7 and 8 is helpful to understand the full pattern of tonal scaling.

2.3 Relational scaling

The third prediction has never been addressed in this form before and the remainder of this chapter is dedicated to its empirical assessment. It posits that the scaling of the f0 value of a pitch accent is essentially relational. Raising or lowering of pitch accents because of information structure only makes sense if it takes place relative to some other pitch accent. The reason for this is that pitch accents are adjusted to register domains which are downstepped relative to their predecessors, and embedded into each other, as shown above. If there is only one prosodic phrase, no downstep and no raising take place, because there is no other register domain relative to which this change can take place. This complex relationship cannot be expressed if pitch accents are addressed directly.

2.3.1 Sentences with fronted objects

In order to test this crucial prediction for the model presented, German sentences with object fronting, as illustrated in (10), are used. In this type of structure, an object is fronted in the sentence-initial, preverbal position. Both a narrow focus on the fronted object and an all-new reading are available.

	H* L	L_I
(10)	[[Die MIETE haben sie wieder mal erhöht]$_F$] $_I$	
	the rent have they again once raised	
	'They have raised the rent again.'	

Figure 9 shows a pitch track of this sentence. The only tonal excursion happens on the fronted object, and the remainder of the sentence has a low and flat intonational contour.[4]

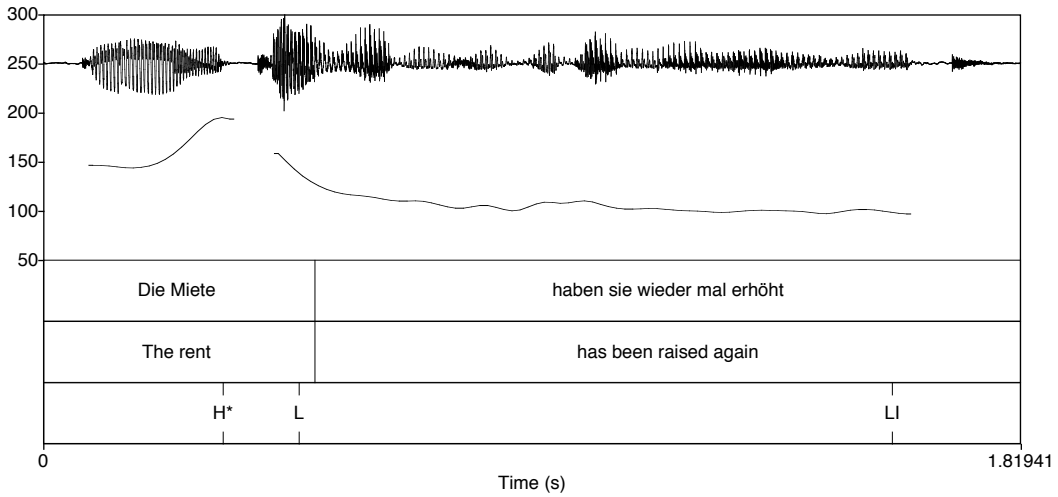


Fig.9 Pitch track of *Die MIETE haben sie wieder mal erhöht*

30 students from the University of Potsdam were recorded. All participants were native speakers of German. Each of them read 12 experimental sentences aloud, as illustrated in (11), as answers to context questions. Additionally, they read 100 unrelated filler sentences presented in a pseudo-randomized order. The object of the target sentences was generic or specific, to check for possible effects of specificity.

[4] See Fanselow (2004) and Fanselow & Lenertová (2008) for syntactic accounts of these sentences.

(11) Wide focus: {Did you go out afterwards?}

Narrow focus: {What did you drink?}

Ein Bier haben wir getrunken./ Ein Jever haben wir getrunken.

a beer/a Jever have we drunk 'We drank a beer/a Jever.'

As predicted, in all sentences (altogether 360 realizations: 12 sentences x 30 subjects), a falling pitch accent was realized on the object and no other accent was present.

There was no difference in f0 value between the narrow and the wide focus realization. All instances of the sentences were realized with a single accent on the object. There were some differences in the average f0 of the objects and the verbs (see Figure 10). In the wide focus condition, the specific objects always had a lower pitch than the generic ones, but the difference is not significant[5] (t = -0.543, df = 54.379, p = 0.5893) and does not relate to the difference in focus context of interest here.[6] Thus no comparison regarding wide or narrow focus was significant (t = -0.1571, df = 693.785, p = 0.8752).

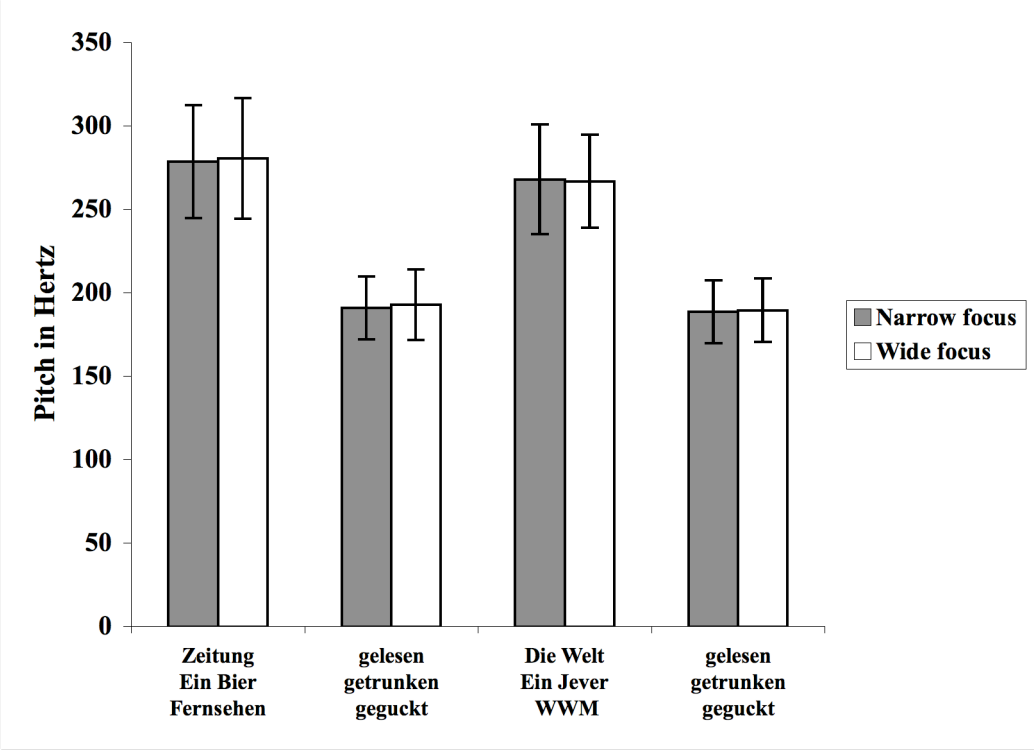


Fig.10 Averaged pitch accents in f0 on the fronted objects of experiment 1. The first and third columns show the object (generic and specific) and the second and fourth columns stand for the verbs.

[5] I am grateful to Heiner Drenhaus for helping me with the statistical analysis of these data. A survey of additional experiments with similar sentences is reported in Féry & Drenhaus (2008).

[6] The remaining comparisons are not significant: verbs in the wide focus condition (t = 1.0112, df = 170.951, p = 0.3134), objects in the narrow focus condition (t = 0.4405, df = 171.677, p = 0.6601), and verbs in the narrow focus condition (t = 0.9323, df = 171.772, p = 0.3525).

No difference in height between an accent on the fronted object in a wide focus context and an accent on the same fronted object in a narrow focus context could be found. This result is compatible with the assumption that there is only one p-phrase, and that, in this case, the height of the top line of the p-phrase does not vary because there is no other register to which the unique p-phrase could adjust. For this reason, the pitch accents are adjusted to a top line which is identical in the wide focus and in the narrow focus conditions.

2.3.2 Subject + verb and object + verb

In a second experiment, sentences consisting of subject + verb or of object + verb were tested. This experiment shows with another very simple syntactic structure that, if there is only one accent, and thus one p-phrase in the relevant VP, no change in f0 value takes place in a narrow focus condition. But as soon as there are two accents in an all-new sentence corresponding to two p-phrases, the height of both accents is affected when a narrow focus is introduced. In such a case, the scaling of the accents is changed. These sentences were again tested in a production experiment, this time with 15 female German students. The experimental sentences are illustrated in (12) to (15). They were inserted in a wide focus (WF) context and in a narrow focus (NF) condition. Altogether, there were 4 conditions, thus a 2 × 2 factorial design, and 6 sentences were constructed for each condition. Altogether 360 realizations were produced.[7]

(12) Subject, WF:

Q: {Why can't I find the ball?}

A: Nun, wahrscheinlich haben ihn [die Kinder mitgenommen]$_F$.
well probably have it$_{ACC}$ the children taken-away
'Well, probably the children took it away.'

(13) Subject, NF:

Q: {Who took the ball away?}

A: Nun, wahrscheinlich haben ihn [die Kinder]$_F$ mitgenommen.

[7] These sentences are part of a larger experiment in which three languages were compared for their register properties: German, Hungarian and Japanese (see Ishihara & Féry 2008).

(14) Object, WF:

Q: {What did the children do?}

A: Nun, wahrscheinlich haben sie [den Ball mitgenommen]$_F$.

well probably have they the ball taken-away

'Well, probably they took the ball away.'

(15) Object, NF:

Q: {What did the children take away?}

A: Nun, wahrscheinlich haben sie [den Ball]$_F$ mitgenommen.

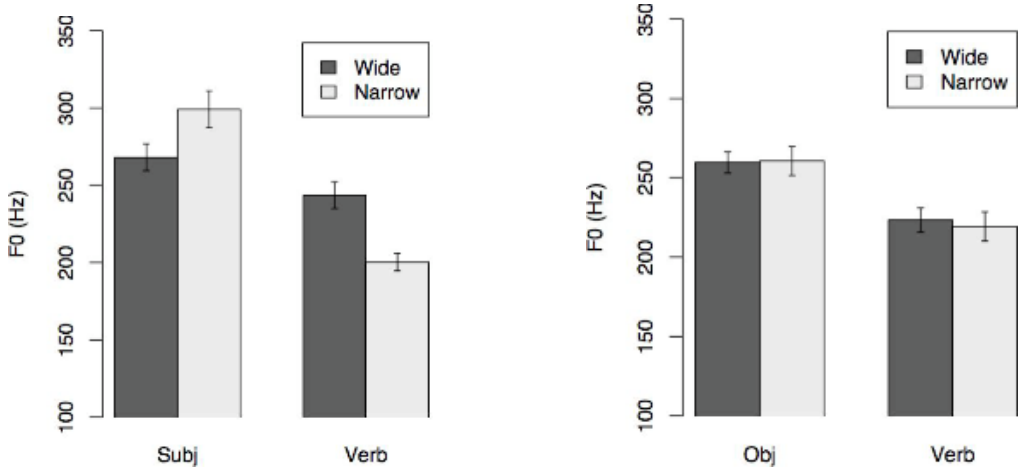
The subject and the object sentences had a different accent pattern. An accent was always realized on the subject or on the object, but (16) shows that an accent was also sometimes realized on the verb in the subject sentences in the wide context condition (51 times in 90 utterances, 57% of the time). Otherwise no accent was produced on the verb, except for one case in an object sentence in the wide focus condition, which can be analyzed as an error.

(16) Realized pitch accents on the verb

Subject sentences in wide focus	51 (57%)
Subject/Object sentences in narrow focus	0 (0%)
Object sentences in wide focus	1 (1%)

Figure 11 shows the pitch height on the verb and on the subject/object in all conditions.

Fig.11 Mean f0-peak on the argument and on the verb (with 95% confidence interval)



Because of the optional accent on the verb in the subject sentences, the verb was scaled significantly higher in the wide focus context than in the narrow focus context. This happened only in the subject sentences. The pitch accent on the verb in 57% of the cases had the effect of considerably raising the average f0 of this constituent.

A secondary effect of the optional accent on the verb was a difference in the height of the pitch accent on the subject, which was higher in the narrow focus condition (when the verb was never accented) than in the wide focus condition (when the verb was sometimes accented).

Turning to the f0 value of the unique accent in the object sentences, there was no significant difference between the pitch accent heights in the narrow and the wide contexts. This result confirms what was observed in the preceding experiment. The difference between the accent pattern and the concomitant f0 value of the constituents is explained by phrasing. The object and the verb form only one p-phrase, both in the all-new and in the narrow focus contexts (Krifka 1984, Jacobs 1993), and, as a consequence, no change occurs in the scaling of the accents.

The subject, in contrast, optionally appears in a separate p-phrase. Following a suggestion by Gisbert Fanselow (p.c.), the subject can syntactically remain in-situ in the VP, which leads to a unique p-phrase, as we saw above for the unaccusative sentences, or the subject may be fronted into the Spec IP position. In this latter case, the subject and the verb are separated in two p-phrases; see Fanselow (2004) and Frey (2004) for 'stylistic' or 'formal' fronting of one constituent in V2 sentences. The difference in phrasing between the object and the subject sentences is illustrated in (17). In (17a,b), the two options for the verb are shown for the subject sentences, and (17c) shows the unique phrasing in the object sentences.

(17) Phrasing in subject and object sentences

Subj: a. $[_{TP}$ Aux O_{pron} S V $]_P$

b. $[_{TP}$ Aux O_{pron} $S_i]_P$ $[_{vP}$ t_i V $]_P$

Obj: c. [Aux S_{pron} O V $]_P$

The difference in phrasing correlates with a difference in metrical structure. Every p-phrase has a head, which means that the two phrasing options for the subject

sentences in (17) correspond to different accent patterns, shown in (18) for the wide focus context. Both (18a) and (18c) have only one p-phrase, and thus one metrical head, but (18b) has two heads.

(18) Metrical pattern in the subject and object sentences in the wide focus context

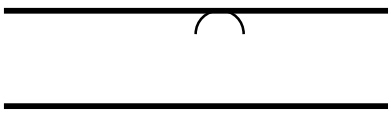
```
      ×
a. ( S    V )P

     ×   ×
b. ( S ) ( V )P

     ×
c. ( O   V )P
```

The phrasing exemplified in (18) also corresponds to different register domains, as shown in Figure 12. (18a) and (18c) have only one p-phrase (Figure 12a). However (18b) has two p-phrases (Figure 12b).

a.



b.

Fig.12 Difference in phrasing between the object and the subject sentences of experiment 2

Consider next what happens in the narrow focus condition, as illustrated in Figure 13. In the conditions (18a) and (18c), shown in Fig. 13a, nothing changes when the unique accent stands for a narrow focus. The register domain corresponding to the unique p-phrase has no other domain to which it can adjust. But in Fig. 13b, narrow focus on the subject raises the top line of the first p-phrase, and lowers the top line

of the second p-phrase, at least in those cases in which the subject is in a different p-phrase from the verb.

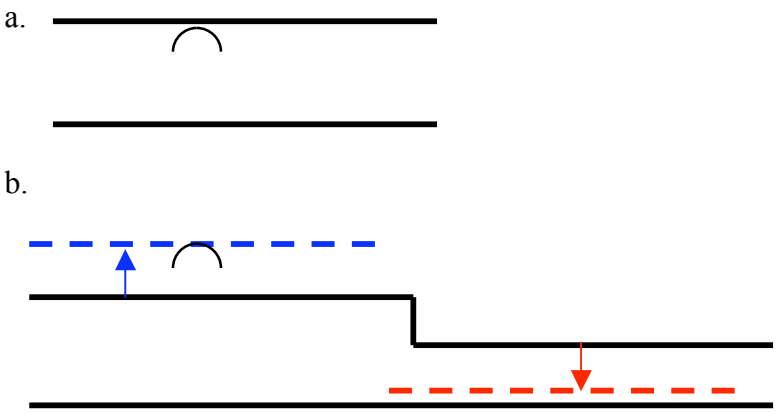


Fig.13 Pitch register change in the two p-phrases environment

The pattern shown in Fig. 13 provides an explanation for the last property of the results in Fig. 11, namely the difference in height between the f0 value of the subject and that of the object in the narrow focus condition. This value is higher on the subject than on the object. In the subject sentences with two p-phrases, a narrow focus has the effect of raising the corresponding top line, as shown in Fig. 13b. Raising the top line is a purely relational effect, in which the register of one p-phrase is changed relative to the register of another p-phrase. In the object sentences and in the remaining cases of subject sentences, no readjustment is needed, since there is only one p-phrase. The verb has no accent in both cases, but the reason for the absence of accent is different. In the object sentences and in part of the subject sentence, this is due to the fact that only one p-phrase is formed on the entire sentence. In half of the subject sentences, the absence of pitch accent is due to compression of the post-nuclear register. These are the cases which relate to an increase of the height of the pitch accent on the subject, as shown in Fig. 13b.

The metrical structure alone cannot account for this difference. It is sometimes assumed that a focus projects a pitch accent at a certain level of the prosodic structure (see, for instance, Selkirk 2002). Then both in (19a) and in (19b), an additional prominence is needed on the argument. Such a representation, however, leads to the expectation that the pitch height of a narrow focus does not depend on the presence of additional p-phrases. The pitch accents of the narrowly focused subject and object are expected to be identical.

(19) Metrical pattern in the subject and object sentences in the narrow focus context

×	×
×	× ×
a. $(S_F/O_F \quad V)_P$	b. $(S_F)(V)_P$

Also, if the metrical structure only reflects the relationship between accents, the difference between object and subject sentences just accounted for is not expected. In both cases, a higher column of beats corresponds to a stronger pitch accent.

3. Discussion and conclusion

In short, the data presented in section 2 confirm that p-phrasing and metrical structure are not sufficient to account for all f0 values observed in declarative sentences in German. A third component is needed which accounts for the relative scaling of accents. This component has been shown to be an abstract modeling of f0 registers corresponding to p-phrases.

Traditionally it has been assumed that pitch accents can change their height on an individual basis: a narrowly focused word is higher than it would be in an all-new context. Similarly a given constituent has a lower f0 value. In other words, accents are changed one by one.

The new aspect introduced in this paper is that the f0 height of pitch accents is interpreted in relationship to neighboring accents. A change in one part of a sentence triggers changes in the other parts of the sentence. Accents are purely relational. This explains why some narrowly focused accents increase their height while others do not. In particular, an accent as the head of a unique p-phrase does not change its value when its information structure is changed. This has been demonstrated with data involving object fronting, as well as with the object sentences of experiment 2. However when an accent standing for a narrow focus is in a sentence containing more than one p-phrase, the scaling of f0 values is modified, and a narrow focus increases the f0 height. This was illustrated with the subject sentences of experiment 2.

References

Beckman, M.E. & J.B. Pierrehumbert (1986) Intonational structure in Japanese and English. *Phonology Yearbook* 3, 255-309.

Berg, Rob van den, Carlos Gussenhoven & Toni Rietveld (1992) Downstep in Dutch: Implications for a model. In Gerhard Docherty & D. Robert Ladd (eds.) *Papers in Laboratory Phonology II: Gesture, segment, prosody.* Cambridge: Cambridge University Press. 335-367.

Bruce, G. (1977) Swedish word accent in sentence perspective. *Travaux de l'Institut de Linguistique de Lund* 12, Gleerup, Lund.

Chomsky, N. & Halle, M. (1968) *The sound pattern of English*. New York: Harper and Row.

Clements, G.N. (1981) The hierarchical representation of tone features. In: Clements, G.N. (ed.) *Harvard studies in phonology* Vol II.

Diesing, M. (1988) Bare plural subjects and the stage/individual contrast. In Krifka, M. (ed) Genericity in language: Proceedings of the 1988 Tübingen Conference. Tübingen: University of Tübingen.

Fanselow, Gisbert (2004). Cyclic Phonology-Syntax-Interaction: Movement to First Position in German. In: Information Structures 1. ISIS. Working Papers of the SFB 632. Potsdam.1-42.

Fanselow, Gisbert & Denisa Lenertová. 2008. Left peripheral focus: Mismatches between syntax and information structure. Ms., University of Potsdam and University of Leipzig.

Féry, Caroline & Heiner Drenhaus (2008) Single Prosodic Phrasing. In: Interdisciplinary Studies in Information Structures 10. Working Papers of the SFB 632. Potsdam. 1-44.

Féry, Caroline and Frank Kügler (2008): German as a tone language. *Journal of Phonetics*.

Féry, Caroline & Vieri Samek-Lodovici (2006) Focus projection and prosodic prominence in nested foci. *Language* 82.1. 131-150.

Féry, Caroline & Hubert Truckenbrodt (2005) Sisterhood and Tonal Scaling. Studia Linguistica. Special Issue "Boundaries in intonational phonology". 59.2/3. 223-243.

Frey, Werner (2004): A medial topic position for German. *Linguistische Berichte* 198. 153-190.

Fuchs, Anna (1976): 'Normaler' und 'kontrastiver' Akzent. *Lingua* 38, 293-312.

Gussenhoven, Carlos (1992): Sentence accents and argument structure. In I. Roca (ed.) *Thematic structure: Its role in grammar*. Berlin & New York: Foris, 79–106.

Gussenhoven, Carlos (2004) *The Phonology of Tone and Intonation*. Cambridge University Press.

Halle, M. & J.-R. Vergnaud (1987) *An essay on stress.* Cambridge: MIT Press.

Ishihara, Shinichiro (2007): Major phrase, focus intonation and multiple spell-out (MaP, FI, MSO). *The Linguistic Review* 24: 137–167.

Ishihara, Shinichiro and Caroline Féry (2008) Syntactic and Prosodic Realization of Information Structure in German, Hungarian and Japanese. Ms. Potsdam.

Ito, Junko, and Armin Mester (2007) *Prosodic Categories and Recursion.* Handout of a talk presented at WPSI3, Indiana University, September 14, 2007.

Jacobs, Joachim (1993). Integration. In *Wortstellung und Informationsstruktur*, Marga Reis (ed.): 64–116. Tubingen: Niemeyer.

Jun, Sun-Ah, Ed. (2005) *Prosodic Typology. The Phonology of Intonation and Phrasing.* Oxford University Press.

Kahnemuyipour, Arsalan (2004). *The syntax of sentential stress*. Unpublished Ph.D dissertation, University of Toronto.

Kratzer, A. (1988) Stage-level and individual-level predicates. In Krifka, Manfred (ed) Genericity in language: Proceedings of the 1988 Tübingen Conference. Tübingen: University of Tübingen.

Kratzer, Angelika & Elizabeth O. Selkirk (2007) Phase theory and prosodic spellout: The case of verbs. *The Linguistic Review* 24: 93–135.

Krifka, Manfred (1984) *Fokus, Topik, syntaktische Struktur und semantische Interpretation*. Ms, München.

Ladd, R. (1990): Metrical representation of pitch register. In Kingston, J. & M. Beckman (eds.) *Papers in laboratory phonology I: between the grammar and physics of speech*. Cambridge: Cambridge University Press, 35–57.

Laniran, Yetunde O. & G.N. Clements (2003) Downstep and high raising: interacting factors in Yoruba tone production. *Journal of Phonetics* 31. 203-250.

Liberman, M.Y. (1978) *The intonational system of English*. Indiana University Linguistic Club, Bloomington.

Liberman, M. & Pierrehumbert, J. (1984) Intonational invariance under changes in pitch range and length. In Aronoff, M. & R.T. Oehrle (eds.) *Language sound structure*. Cambridge: MIT Press, 157-233.

Liberman, M. & Prince, A. (1977) On stress and linguistic rhythm. *Linguistic Inquiry* 8, 249-336.

McCarthy, John J. & Alan S. Prince (1990) Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory* 8: 209-283.

Nespor, M. & I. Vogel (1986): *Prosodic phonology*. Dordrecht: Foris.

Pierrehumbert, J. (1980): *The phonology and phonetics of English intonation.* PhD dissertation, MIT.

Poser, William J. (1984) *The phonetics and phonology of tone and intonation in Japanese.* PhD dissertation. MIT, Cambridge, Massachussetts.

Selkirk, E. (1984): *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge: MIT Press.

Selkirk, E. (1986): On derived domains in sentence phonology. *Phonology Yearbook* 3: 371–405.

Selkirk, Elisabeth O. (2000) The Interaction of Constraints on Prosodic Phrasing. In: Horne, Merle (ed.) *Prosody: Theory and Experiment.* Dordrecht. Kluwer. 231-261.

Selkirk, Elisabeth. (2002). Contrastive FOCUS vs. presentational focus: Prosodic Evidence from Right Node Raising in English. In *Speech Prosody 2002: Proceedings of the First International Speech Prosody Conference,* B. Bel and I. Marlin (eds.). Aix-en-Provence: Laboratoire Parole et Langage, 643–646.

Truckenbrodt, Hubert (1999) On the Relation between Syntactic Phrases and Phonological Phrases. *Linguistic Inquiry* 30:2. 219-255.

Truckenbrodt, Hubert (2004) Final lowering in non-final position. *Journal of Phonetics* 32. 313-348.

Truckenbrodt, Hubert (2007) Upstep on edge tones and on nuclear accents. In Carlos Gussenhoven and Tomas Riad (eds.) *Tones and tunes: Studies in word and sentence prosody*. Vol 2. Berlin: Mouton. 349-386.

Wagner, Michael. 2005. *Prosody and recursion.* Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Mass.

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics*, 27: 55-105.