Recursion in Prosodic Structure*

Caroline Féry
Goethe University Frankfurt

ABSTRACT. This paper examines some aspects of the prosody of German which speaks for a recursive organization. This is done with the help of results of production experiments which especially addressed this issue. Tone scaling was examined, and it could be shown that the pattern of downstep and reset follows the recursive morpho-syntactic structure with exactitude.

Keywords: prosody, intonation, recursion

1. Introduction

This paper examines the hypothesis that higher prosodic constituents are recursive. If this hypothesis is correct, as we believe it is, it implies important departures from the standard literature on prosody-syntax interface. First, only part of the Strict Layer Hypothesis (Nespor & Vogel (1986), Selkirk (1984)) is correct in its assumptions, namely that prosodic constituents of level n cannot dominate larger constituents n+1 or more generally n + m (m > n). But recursion implies that prosodic constituents of level n may dominate other constituents of level n, which are entirely contained in the larger ones. Second, prosody is organized at least in part like syntax, and is not, as previously envisaged, completely flat. The novel claim of this paper is that a recursive prosodic structure is always mapped to a recursive morpho-syntactic structure.

The assumptions of Strict Layer Hypothesis could be taken for granted as long as no recursive structure was found in any language. But recursive prosodic constituents have been discovered very early in a number of languages. Ladd (1990, 1996) claimed that intonation phrases are recursive in English. He called them 'compound domains' to emphasize the observation that compound phrases behave like recursive compound nouns in forming larger prosodic constituents consisting of smaller ones of the same kind. Kubozono (1989) proposed that Major Phrases have a binary structure, following the recursive structure of syntax. Ishihara (2003, 2004, 2007) examined the prosodic structure of embedded wh-phrases in Japanese, as compared to sentences in which the embedded wh-word takes its scope in the root clause, and came to the conclusion that larger prosodic phrases are recursive (see also Deguchi & Kitagawa (2002) and Shinya, Selkirk and Kawahara (2004), among others). Kubozono (2007) examined the influence of focus on wh-words, which seems to elicit a 'new' prosodic domain. He concluded that recursion explains the results better than assuming a new prosodic domain. Ito & Mester (2008) were interested in complex prosodic words in Japanese. They showed that compound words need four different patterns involving Prosodic Words and Phonological Phrases. Some of the structures show recursion. Ito & Mester (2010)

generalized on prosodic structure and showed that the assumption of recursive prosodic domains can only be advantageous for the prosodic theory in general. In sum, quite an impressive literature has shown that Japanese needs recursion of its higher prosodic domains, namely the ones which are mapped to morpho-syntactic structures. See also Selkirk's (2009) recent work for a review of the literature on recursion in Japanese and a novel approach.

In the following, I would like to review some of the evidence from German which show that, in this language as well, recursive prosodic structure of the higher prosodic domains is the rule rather than the exception.

2. Prosodic Structure

The standard Prosodic Hierarchy, as proposed by Nespor & Vogel (1986) and Selkirk (1984), and many others after them, is organized as in (1).

(1)	U		Utterance
	IP	(ι)	Intonation phrase
	PPh	(φ)	Phonological Phrase (= p-phrase)
	PW	(ω)	Prosodic Word
	F		Foot
	σ		Syllable
	μ		Mora

The Strict Layer Hypothesis in (2) formulates several restrictions at once (Nespor & Vogel (1986), Selkirk (1984:26)). First, that constituents of level n dominating constituents of a lower level may not skip any level. The only constituents which they may dominate are of level n-1. And second, that constituents of level n may not dominate constituents of the same level (no recursivity), or constituents of a higher level.

(2) Strict Layer Hypothesis (Nespor & Vogel (1986:7))

Principle 1. A given nonterminal unit of the prosodic hierarchy Xp is composed of one or more units of the immediately lower category, Xp-1.

Ideally, a sentence is parsed exhaustively at all levels of the prosodic hierarchy, and universal prosodic domains are formed according to principles. An example of exhaustive parsing is given in (3) for English.

```
(                                                        ×  )   IP
(       ×    ) (               ×  )   (       ×     ) (      ×  )   p-phrase
( × )  (×    ) (×  ) (       ×    )   (       ×     ) (×    )( ×  )   PW
   F      F        F    F    F            F    F        F      F
   σ      σ  σ    σ   σ   σσ   σ        σ   σσ   σ   σ     σ σ
  μμ    μμ μμ   μμ μ   μ μμ   μμ      μμ μ μμ μμ μμ    μμ μμ
```

(3) Miss Martin went to the market with a basket full of eggs

But this perfect layering breaks down in a number of circumstances. First, a prosodic word is not always completely parsed in feet. In (4), the first two syllables build a trochaic foot together, but the last syllable is too light to form a foot on its own. Thus, it must be the

case that a prosodic word may dominate a syllable (Hayes (1982), Liberman & Prince (1977)). The same is true for German (in fact (4) may be a German word).

$_{F}(\times .)$

(4) Minima

Second, prosodic words may be recursive. The German compound word in (5) consists of four individual words put together.

(5) [[[Bundes-gesetz]entwurfs]komittee]
federal law draft committee

Third, there is some uncertainty as to the levels between prosodic words and prosodic phrases. Some function words may not be said to form prosodic words by themselves, because they do not fulfill the weight conditions for a prosodic word. Nespor & Vogel (1986) try to solve the problem by assuming an intermediate level, the 'clitic group.' But the clitic group can only take care of part of the problematic material.

As we will see below, the question of recursion arises in relation with higher prosodic domains, like prosodic phrases (p-phrases) and intonation phrases (i-phrases). We will see that assuming recursions of these domains solves some layering problems.

3. What is Prosodic Recursion?

Hauser, Chomsky & Fitch (2002) define recursion as the basic operation that allows the generation of a potentially infinite array of discrete expressions out of a finite set of elements. The set of finite elements is hierarchically organized.

For prosody, recursion implies a set of prosodic domains which can be repeated at each level of the hierarchy. We already saw that lower prosodic domains cannot dominate higher ones. Either the domains are repeated linearly, or they are contained within each other. The former method is known as iteration, and is universally admitted in the literature on prosodic structure. It is illustrated in (6) with a list, see for instance Nespor & Vogel (1986), Liberman & Pierrehumbert (1984), van Heuven (2004) for the prosodic realization of lists. In an iterative structure, as in (6), the prosodic domains iterate but do not overlap.

(6) (Anna made some errands and bought) [a bottle of orange juice]$_{P}$, [an apple]$_{P}$, [sugar]$_{P}$, [butter]$_{P}$, [a pair of socks]$_{P}$

But here, the second meaning of recursion will be investigated: a prosodic domain of level n may be contained in another, larger domain of the same type n. We thus make a principled distinction between *iteration* of prosodic domains n, and *recursion* of prosodic domains n (see also Hunyadi (2006) for this distinction). In recursion proper, a center-embedded clause occurs in the middle of a main clause, which, as a result, is divided into two parts. In such structures, a single syntactic constituent is separated by another constituent of the same kind. English has center-embedded relative clauses and sentential complements which can be called recursive. An example appears in (7).

(7) [The girl [who wears a white shirt] wants to become a lawyer]

The alternative implies that the root clause is parsed into two different prosodic

domains, one containing the subject, and another one containing the VP, so that the entire sentences presents iteration of prosodic domains. The question arises as to the criteria used to decide between the two options. Ito & Mester (2008) use syllabification and accent placement as diagnostic for recursive structure in prosody which are used differently in different kinds of compounding. Ladd (1990) uses pitch range, and more specifically upstep, as a test for prosodic recursion at higher levels of prosodic structure. This criterion was also used by a number of authors, like Ishihara (2003), Kubozono (1989, 2007) and Féry & Kentner (2010). Hunyadi (2006) and Wagner (2005) measured variation in duration. The common observation of all researchers is that some recursion of prosodic domains has to be allowed at least at some levels of the hierarchy.

4. Recursion at the Level of the Intonation Phrase

In Féry & Truckenbrodt (2005), we investigated for German whether a sequence of three syntactically and semantically related English sentences are in a downstep and/or reset relationship, depending on what their internal syntactic and prosodic structure looks like. The design of this experiment was close to the one conducted by Ladd (1990) for English. In Féry & Truckenbrodt, two conditions were examined in a production experiment with the patterns shown in (8) and (9).

(8) 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 possesses a Lada, and the violinist drives a Wartburg.'

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

{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 possesses a Lada, and the violinist drives a Wartburg, while the painter has a Jaguar.'

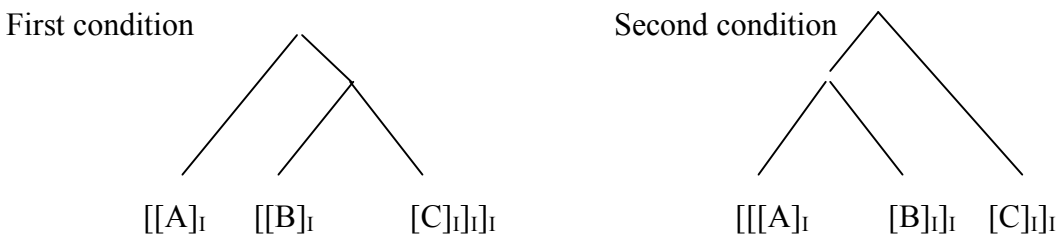


Fig.1 Two conditions in the experiment reported in Féry & Truckenbrodt (2005)

The difference between the prosodic structure of the two conditions is illustrated in Figure 1. 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 scaling relationship with 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 an i-phrase as well.[1] The tonal structure of a sentence of the first condition is shown in (10). Important for the pitch scaling are the values of the first H tone in each sentence.

(10) {Why does Anna think that sportsmen have more expensive cars than crafstmen?}

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]$_P$]$_I$]$_I$

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

A production experiment was conducted in Potsdam with five students, native speakers of Standard German, who uttered 32 experimental sentences each. The pattern which emerged from the experiment was that the first condition had a downstep pattern throughout, as shown in Figure 2, but the second condition elicited a reset on the C sentence, as shown in Figure 3. The first high tone of this sentence was slightly higher than the first tone of sentence B. Moreover, this tone was much higher than it was in the first condition.

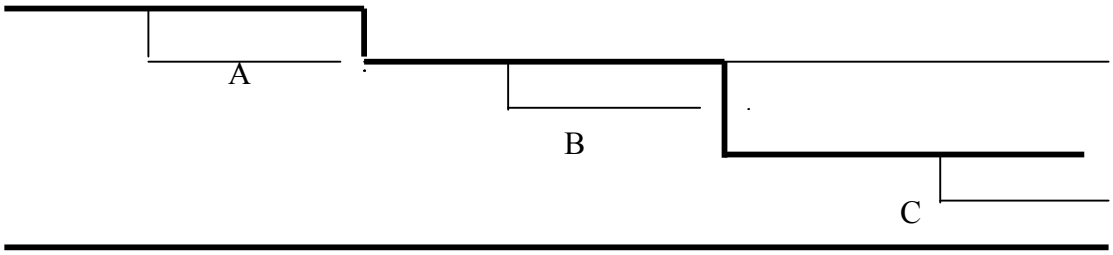


Fig.2 Results of Féry & Truckenbrodt (2005) for the first condition

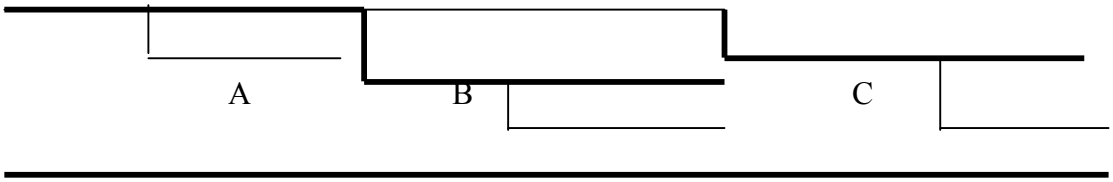


Fig.3 Results of Féry & Truckenbrodt (2005) for the second condition

This result speaks for recursion proper rather than for iteration of the i-phrases. The tone scaling is sensitive to the syntactic structure of the utterance. Downstep and reset are both playing 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.

5. Recursion at the Level of the P-Phrase

In this section, recursion at the level of the p-phrase is examined for German. First, a semi-spontaneous sentence from Féry, Hörnig & Pahaut (2010) is used to present the kind of evidence that speaks for recursion. In a second step, results of an experiment are reported.

First, let us examine the sentence reproduced in (11) and illustrated in Figure 4. It was uttered in the framework of an experiment asking participants to describe the location of three toy animals placed on a table in front of them.

(11) L*H_P L* H_P H* H*L_P

[[[Also jetzt steht]$_P$ [[links]$_P$ [der Gorilla]$_P$]$_P$ [[rechts]$_P$ [neben dem Gorilla]$_P$]$_P$

so now stands left the gorilla right beside the gorilla

L*H_P L*H_P H*L_P H*L_I

[das Pferd]$_P$]$_P$ [[und neben dem Pferd]$_P$ [rechts]$_P$ [der Löwe]$_P$]$_P$]$_I$

the horse and beside the horse right the lion

'So now the gorilla is standing to the left; the horse is standing to the right beside the gorilla; and the lion is standing beside the horse to the right.'

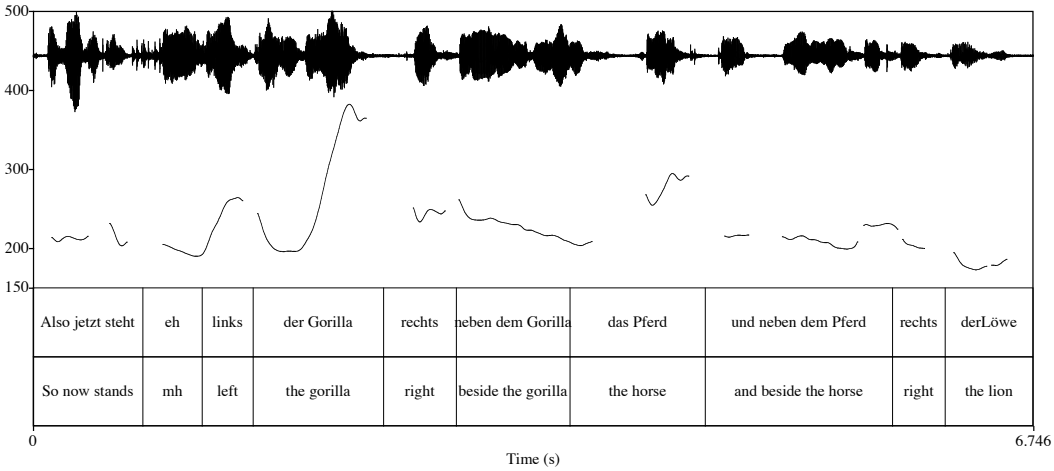
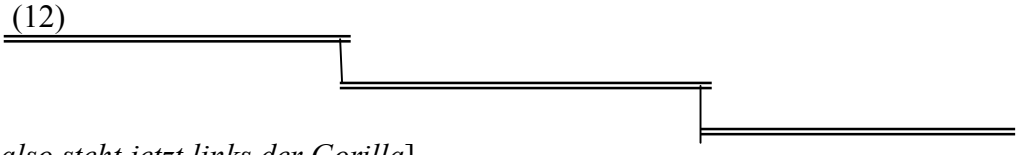


Fig. 4. A spontaneous sentence with a complex pitch scaling

Each of the localizations of an animal is a p-phrase. As a result, there are three p-phrases at the highest level of p-phrasing, in a clear downstep relationship to each other. Together, they are organized in an i-phrase. This is shown in (12).

(12)



[*also steht jetzt links der Gorilla*]$_{P1}$

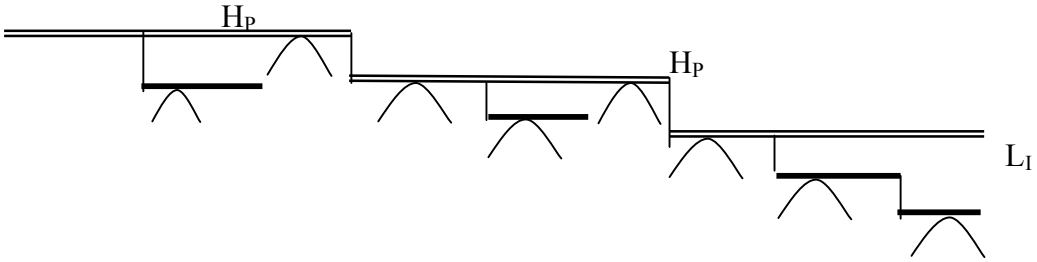
[*rechts neben dem Gorilla das Pferd*]$_{P2}$

[*und neben dem Pferd rechts der Löwe*]$_{P3}$

Each of these p-phrases contains embedded phrases. First, the animal to be localized (called *locata*) *der Gorilla*$_{Loc1}$, *das Pferd*$_{Loc2}$ and *der Löwe*$_{Loc3}$, which are in a downstep relationship to each other, the highest pitch values being 388, 303 and 178 Hz, respectively. Second, the locative expressions: *links*$_{Lx1}$, *rechts neben dem Gorilla*$_{Lx2}$ and *und neben dem*

Pferd rechts$_{Lx3}$ (highest pitch values 269, 253, and 230 Hz, respectively). Inside the second complex locative expression, we have another downstep, *rechts* being slightly higher than *neben dem Gorilla* (253, 235 Hz). In the third one, *neben dem Pferd* is slightly higher than *rechts* (230, 204 Hz).

The first two highest-level p-phrases end at a high level. The reason is that the animals are strongly contrasted with each other and cause a focus boost in their respective domains. While this is true for *Gorilla* and *Pferd*, it is not the case for *Löwe*, which is final and attracted by the final low tone. For this reason, it is only subject to downstep. The last high tone in a p-phrase is scaled at the level of the reference line of its own p-phrase, as illustrated in (13), except for the final one.



(13)

The metrical grid for the sentence is given in (14). Every pitch accent is the head of its p-phrase. At the highest level of phrasing, the heads are strongest, and they become weaker when the p-phrases they head are embedded. A deeper embedding goes together with a weaker metrical position.

(14)

		x					x						x		
	x		x			x			x			x			
[[	[x]		[x]]		[[[	x]	[x]]	[x]]		[[[	x]	[x]]		[x]]]	
	269		388			253	235	303			230	204		178	

Let us now turn to the second type of experimental results, in the framework of an experiment addressing this question of recursion in a specific way, with the production of coordinated names, like *Anna and Bill or Mary*. Such sequences form an ambiguous structure, in the same way as an arithmetic procedure like *3 minus 2 plus 1*, which can be resolved as 2 or as 0, depending on the order of the operations. Researchers have examined how different groupings of coordinated names or numbers are realized prosodically (as for instance Lehiste (1973), Ladd (1996), and Wagner (2005) for English, Hunyadi (2006) for Hungarian). All authors focus on differences in duration at conjunct boundaries and find a strong dependency between duration of constituents and their place in the coordination structure.

In Féry & Kentner (2010), we investigated pitch scaling between the different elements of coordination, like those shown in (15). *Oder* 'or' and *und* 'and' were used as connectors between names. The former one outside, and the latter one inside the parentheses.

(15) Suse oder Nino oder Mila oder Anna
Suse oder Nino oder (Mila und Anna)
(Suse und Nino) oder Mila oder Anna
Suse oder (Nino und (Mila und Anna))

((Suse und Nino) und Mila) oder Anna
(Suse und Nino) oder (Mila und Anna)

For each item, a context question, spoken by a female native speaker of German, had been previously recorded. The contexts were presented together with a target sentence both visually on screen and aurally over headphones. To emphasize the structure of the target sentence, it was displayed with parentheses, as in (15). 21 female participants, all native German speakers from the Berlin area (North Germany), read out the complete set of target sentences (n=24) interspersed with numerous fillers.

The full set of 504 target sentences was hand-annotated by two phonetically trained students and subjected to phonetic analysis. Durations of each name plus the following pause were measured. And the F0 minima and maxima of each name and connectors were measured as well. The results are shown in Figure 5.

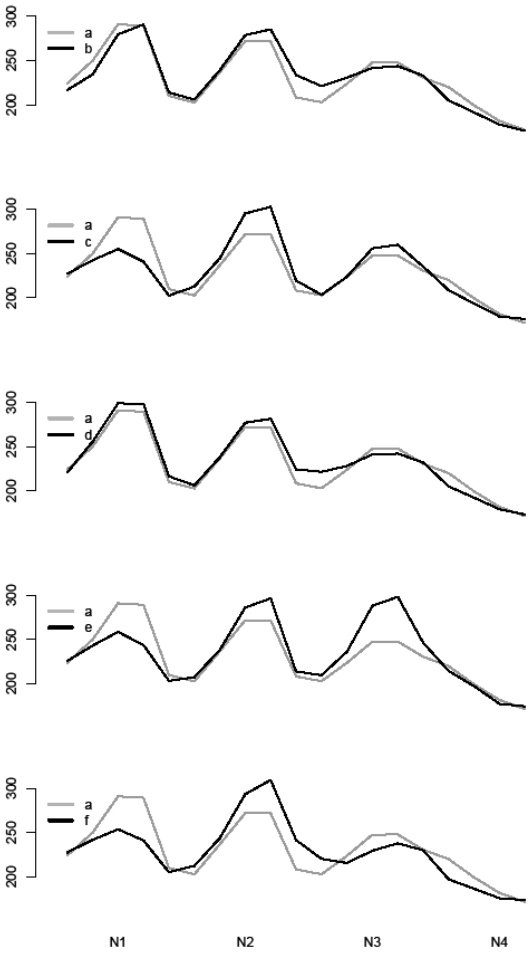


Fig.5. F0 tracks (in Hz) of baseline (grey) plotted against other conditions (black).

The results show that prosodic structure reflects grouping and embedding in a precise way. The lower pitch and shorter duration observed on the left member of groupings show that this constituent is reduced relatively to baseline, the realization without grouping. A boundary between two constituents has the effect of strengthening the element before the parenthesis. It could also be shown that simplex elements in an expression containing groupings have increased duration and higher pitch to achieve similar prosody to complex

elements at the same level of syntactic embedding. As a result, German interprets syntactic structure with exactitude. The experiment demonstrates that, at least in this language, prosody as a whole supports the rendition of syntactic structure. For our aims, it can also be shown that only a recursive prosodic structure, following the syntactic structure, can account for the differences in the renditions of the patterns in (15).

6. Conclusion

In this paper, evidence for a recursive prosodic structure has been advanced for German. For reason of space, the paper has concentrated on only two prosodic domains, and has evaluated only some experimental evidence speaking for recursion in prosody. As has been shown with the help of production experiments, recursion is present at the higher levels of prosodic domains, where morpho-syntax is mapped to prosody.

It is in fact surprising that prosody has been assumed to have a flat structure, with a strict layering and no recursion, for such a long time. The reason is to be found in the understudy of prosody, and the concentration on very simple syntactic structures.

****

Notes

[*] I would like to thank the Phonological Society of Japan for inviting me to give a talk in August 2009, and for giving me the opportunity to write this paper. I profited greatly from the comments of the audience. Many thanks also to my colleagues at the University of Potsdam.

[1] This assumption differs from the pattern in Féry & Truckenbrodt (2005), in which we were more traditional in avoiding recursion of prosodic domains.

References

Deguchi, Masanori and Yoshihisa Kitagawa (2002) "Prosody and Wh-questions," *Proceedings of the 32nd Annual Meeting of the Northeastern Linguistics Society*, ed. by M. Hirotani, 73–92. Amherst MA: GLSA, UMass Amherst.

Féry, Caroline, Robin Hörnig & Serge Pahaut (2010) Phrasing in French and German: an experiment with semi-spontaneous speech In: Gabriel, Christoph & Conxita Lleó (eds.) *Intonational Phrasing at the Interfaces: Cross-Linguistic and Bilingual Studies in Romance and Germanic*. Amsterdam: Benjamins.

Féry, Caroline and Gerrit Kentner (2010) *The Prosody of Embedded Coordinations in German and Hindi*. In Proceedings of Speech Prosody. Chicago.

Féry, Caroline and Hubert Truckenbrodt (2005) "Sisterhood and Tonal Scaling," *Studia Linguistica. Special Issue "Boundaries in Intonational Phonology"* 59:2/3, 223-243.

Hauser, Marc D., Noam Chomsky and W. Tecumseh Fitch (2002) "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?" *SCIENCE* 298, 2769-79.

Hayes, Bruce (1982) "Extrametricality and English Stress," *Linguistic Inquiry* 13:2, 227-276.

Hunyadi, László (2006) "Grouping, the Cognitive Basis of Recursion in Language," Kossuth Egyetemi Kiadó,

Debrecen. *Argumentum* 2, 67-114.

Ishihara, Shinichiro (2003) *Intonation and Interface Conditions*. Doctoral Dissertation, Massachusetts Institute of Technology.

Ishihara, Shinichiro (2004) "Prosody by Phase: Evidence from Focus. Intonation–Wh-Scope Correspondence in Japanese," *Interdisciplinary Studies in Information Structures 1. Working Papers of the SFB 632*, ed. by S. Ishihara, M. Schmitz and A. Schwarz, 77-119. Potsdam: University of Potsdam Press.

Ishihara, Shinichiro (2007) "Major Phrase, Focus Intonation and Multiple Spell-Out (MaP, FI, MSO)," *The Linguistic Review* 24, 137–167.

Ito, Junko and Armin Mester (2010) "Recursive Prosodic Phrasing in Japanese," *Prosody Matters: Essays in Honor of Elisabeth Selkirk*, ed. by T. Borowsky, S. Kawahara, T. Shinya and M. Sugahara. London: Equinox Publishers.

Ito, Junko and Armin Mester (2008) "The extended prosodic word," *Phonological Domains: Universals and Deviations*, ed. by J. Grijzenhout, and B. Kabak. Berlin: Mouton de Gruyter.

Kubozono, Haruo (1989) "Syntactic and Rhythmic Effects on Downstep in Japanese," *Phonology* 6, 39-67.

Kubozono, Haruo (2007) "Focus and Intonation in Japanese," *Interdisciplinary Studies on Information Structure 9: Proceedings of the 2nd Workshop on Prosody, Syntax and Information Structure*, ed. by S. Ishihara, 1-27. Potsdam: University of Potsdam Press.

Ladd, D. Robert (1990) "Metrical Representation of Pitch Register," *Papers in Laboratory Phonology 1: Between the Grammar and the Physics of Speech*, ed. by J. Kingston, and M. Beckman, 35–57. Cambridge: Cambridge University Press.

Ladd, D. Robert (1996) *Intonational Phonology*. Cambridge: Cambridge University Press.

Nespor, Marina and Irene Vogel (1986) *Prosodic Phonology*. Dordrecht: Foris.

Lehiste, Ilse (1973) "Phonetic Disambiguation of Syntactic Ambiguity," *Glossa* 7, 107-122.

Liberman, Mark and Janet Pierrehumbert (1984) "Intonational Invariance under Changes in Pitch Range and Length," *Language Sound Structure,* ed. by M. Aronoff and R. T. Oehrle, 157-233. Cambridge: MIT Press.

Liberman, Mark and Alan Prince (1977) "On stress and Linguistic Rhythm," *Linguistic Inquiry* 8, 249-336.

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

Selkirk, Elisabeth (2009) "On Clause and Intonational Phrase in Japanese: The Syntactic Grounding of Prosodic Constituent Structure," *Gengo Kenkyu* 136.

Shinya, Takahito, Elisabeth Selkirk, and Shigeto Kawahara (2004) "Rhythmic Boost and Recursive Minor Phrase in Japanese," *Proceedings of the Second International Conference on Speech Prosody*, ed. by B. Bel and I. Marlien, 183-186. Nara, Japan.

van Heuven, Vincent J. (2004) "Planning in Speech Melody: Production and Perception of Downstep in Dutch," *On Speech and Language: Studies for Sieb G. Nooteboom.*, ed. by H. Quené and V. van Heuven, 83-93. Utrecht University: LOT Occasional Series.

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