Hierarchical prosodic structures in the intonation of center-embedded relative clauses[1]

CAROLINE FÉRY AND FABIAN SCHUBÖ

Abstract

There is no doubt that recursion exists in syntax, but whether this is reflected in prosody is still an open question. In this paper, the prosody of sentences with syntactically recursive center-embedded clauses is examined empirically in two languages. On the basis of the results, we argue for recursion of higher prosodic domains in German. Evidence for this proposal comes from the F_0 range of the matrix clause, which is larger if embedded material is present. It is both higher in the high regions and lower in the low regions than in matrix clauses without embedded material. A second piece of evidence relates to pitch scaling of high tones at the boundaries between clauses, which is sensitive to syntactic embedding. In Hindi, by contrast, no evidence for embedding of prosodic phrases could be found. Prosodic phrases are concatenated and are in a downstep relationship to each other, regardless of the syntactic structure of the sentence. The difference between the prosodic structure of German and Hindi is tentatively explained by their different intonational properties: German is an intonation language with lexical stress and Hindi is a phrase language without lexical stress, in which tonal structures are phrase-based.

1. This paper is part of the project Prosody in Parsing, financed by the DFG. Many thanks to Umesh Patil, who constructed the Hindi sentences and who also performed the recordings in Delhi. Thanks are also due to Caroline Magister, Verena Thießen and Georg Höhn for technical assistance, as well as to Anja Arnhold and Gerrit Kentner for helpful comments. We are grateful to two anonymous reviewers, whose careful and helpful comments led to improvements in the paper. Last but not least, Kirsten Brock checked the last version of this paper and corrected English.

The Linguistic Review 27 (2010), 293–317
DOI 10.1515/tlir.2010.011

0167–6318/10/027-0293

1. Introduction

1.1. What is recursion?

This paper investigates the prosody of main clauses accompanied by center-embedded relative clauses in German and Hindi. The correlate of prosody used for this investigation is pitch scaling. The results of the experiments presented below are intended to shed some light on the issue of recursion at the interface between syntax and prosody (see Ladd 1986, 1990, Wagner 2005, Ito and Mester 2008, to appear, and Hunyadi 2006 among others).

Hauser, Chomsky and 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 means that each level of the hierarchy of prosodic domains can be repeated. Either the domains are repeated linearly, or they are contained within each other. The first method is known as iteration, and is universally accepted in the literature on prosodic structure. It is illustrated in (1) with a list; see for instance Nespor and Vogel (1986) and Liberman and Pierrehumbert (1984), for the prosodic realization of lists. In such an iterative structure, the prosodic domains iterate but do not overlap. Downstep applies across a sequence of p-phrases in German, as in English. In (1), the whole sentence is an intonation phrase and every element of the list is in a separate phrase. The highest tone of every p-phrase is on a lower level than the highest tone of the preceding p-phrase. We claim that this is indicative of iteration of prosodic domains at the same level of embedding (see Pierrehumbert and Beckman 1988 for English, Féry 1993 for German).

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

We are interested in the second meaning of recursion in this paper: 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 Hunyadi 2006 for this distinction). In recursion proper, a center-embedded clause, or even two in the case of (2a), occurs in the middle of a main clause, which, as a result, is separated into two parts. In such structures, a single syntactic constituent is divided up by another constituent of the same kind. German has center-embedded relative clauses and sentential complements which can be called recursive in the sense used here.

(2) a.

[$_{subj}$	*Die*	*Jungen,*	[$_{CP1}$	*die*	*das*	*Kind,*	[$_{CP2}$	*das*	*Angst*
	the	boys		who	the	child		who	fear

hat], *ärgern]*, *sind böse*]
has tease are nasty
'The boys who tease the child who is scared are nasty.'

b. *Der Direktor hat seine Ansicht, dass die Firma zu*
the director has his opinion that the firm too
große Verluste einbüßt, der Gewerkschaft
large losses suffers the-dat workers' union
mitgeteilt.
communicated
'The director has communicated to the workers' union his opinion that the firm is suffering too many losses.'

Syntax is unable to restrict the number of embedded clauses that a sentence may have, and prosody seems to be unconstrained in principle as well, as it cannot limit the number of iterated prosodic domains of a sentence. If every element of the list in (1) is a prosodic phrase, it is always possible to add an element and thus add a prosodic domain. Embedded structures such as in (2) might be unlimited in the same way, but in practice the embedding is restricted by understandability. More than two embedded clauses are very difficult to process, both in production and in comprehension.

The central question addressed in this paper concerns recursion of prosodic domains, and especially of higher domains, corresponding to syntactic clauses. If embedded syntactic clauses may be recursive, are the corresponding prosodic domains also recursive, or do they show prosodic flatness?

1.2. Hypotheses about prosodic structure and acoustic correlates

In the experiments presented below, sentences containing two center-embedded relative clauses like (2a) were investigated for their prosody.

It is sometimes taken for granted in the literature that intonational phrases and prosodic structure in general contain sequences of prosodic domains that do not overlap, as required by the Strict Layer Hypothesis (see Nespor and Vogel 1986: 7, Selkirk 1984: 26). This implies that the syntactic center-embedding of (2) could not be reflected in the prosody. In this case, the prosodic structure of a sentence like (2a) is as in (3), and thus does not differ from an iterative structure. If this is confirmed, downstep should take place in a recursive fashion, as explained for (1). Subscripted P in the examples stands for 'prosodic phrase' (or p-phrase), which we assume to be the only prosodic domain between the prosodic word and the intonation phrase. The important aspect here is that, in (3), there is no level of prosodic phrasing at which the prosodic domains are embedded into each other. More formally, there is no level n of prosodic

structure such that a prosodic domain n can dominate another prosodic domain n.

(3) [*Die Jungen*]$_{\mathrm{P}}$ [*die das Kind*]$_{\mathrm{P}}$, [*das Angst hat*]$_{\mathrm{P}}$, [*ärgern*]$_{\mathrm{P}}$ [*sind böse*]$_{\mathrm{P}}$

There is an alternative to (3) in which prosodic structure involves recursion. In this case, a prosodic domain of a certain level n can include another prosodic domain of the same level n (see Ladd 1986 for an early proposal of recursion of Intonation Phrases, and Ito and Mester to appear for a more recent one). In such an approach, the prosodic structure of (2a) is as illustrated in (4).

(4) [*Die Jungen* [*die das Kind,* [*das Angst hat*]$_{\mathrm{P}}$, *ärgern*]$_{\mathrm{P}}$ *sind böse*]$_{\mathrm{P}}$

If we can show that prosody is recursive, then the Strict Layer Hypothesis is only correct up to a certain point. It is correct in requiring a hierarchical organization of prosodic domains, such that a constituent of category p can only be dominated by a higher category p + n. But it is not correct in excluding that a unit of category p can be dominated by another unit p.

The standard prosodic hierarchy model predicts only a few levels of strictly layered prosodic constituents: at the higher levels, which concern us in this paper, these are the Phonological Phrase (ϕ), the Intermediate Phrase (ip) and the Intonation Phrase (IP). A sequence of sentences that are prosodically coherent forms an Utterance (U). These levels are claimed to be sufficient to account for all possible prosodic structures in all languages. However, some considerations have led a number of researchers to doubt the validity of the assumptions made by the Strict Layer Hypothesis (see Wagner 2005, Ito and Mester 2008, Féry and Kentner 2010), and this on the following grounds: First, in complex structures involving numerous phrases, like in name groupings as illustrated in (5), more than three levels (or four if one adds the Utterance) are needed. If every name (indicated here with a letter) is a Prosodic Word and the whole sentence is an IP, we have maximally two levels of prosodic phrasing at our disposal, but at least four are needed. Second, it is extremely difficult or even impossible to find invariant phonetic correlates for every level (see Ito and Mester to appear). It looks rather as if phrasing is relative, in the sense that more or fewer correlates of phrasing are needed in dependence on another phrasing in the same sentence (see also Watson and Gibson 2004 for a similar result in psycholinguistics).

(5) ((A and B) or (C and (D and E) or (F, G and (H and (I or J))))) are invited to my party.

Turning now to the predictions for the tonal part of the prosodic structure, there is a widespread assumption in the literature that the second part of a matrix clause interrupted by a non-restrictive relative clause, a sentential com-

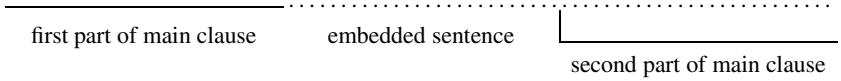

Figure 1: Assumption 1: Invisibility of the parenthetical expressions (recursive prosody)

plement or a parenthetical expression resumes the F_0 declination after the interruption at the same melodic height as it was before the interruption (Cooper and Sorensen 1981, Kutik et al. 1983, Ladd 1986 for English, and Hunyadi 2006 for Hungarian). Let us call this "Assumption 1," which will be tested for German for relative clauses. Cooper and Sorenson (1981) found for sentences like (6a) that the declination of the matrix clause was the same when the parenthetical was present as when it was not (see Figure 1). Kutik et al. (1983) were able to reproduce the results. The sentence (6b) comes from Ladd (1986).

Hunyadi (2006) claims that the same effect appears in Hungarian. He uses not only parenthetical expressions, but also center-embedded relative clauses of the same kind as those we use for German.

In Figure 1, the dotted line stands for the reference line (or top line) of the speaker's register in the first part of the sentence. It continues until the end of the sentence, and is not interrupted by the embedded sentence, which may or may not be at the same level of reference as the matrix sentence. As a result, the second part of the matrix clause is scaled relatively to the first part. A moderate amount of downstep applies between the first and the second part of the matrix clause.

(6) a. *The book on the table, it seems to me, was a gift from my mother.*
b. *My brother, who is a geologist, lives in Denver.*

A second possible pattern is illustrated in Figure 2. It implies that the final part of the main clause is scaled to the embedded clause. In this case, the reference line of the first part of the matrix clause does not serve as the reference for the second part. Instead, the downstepped reference line of the embedded sentence takes over this role. Such a result is more in line with an iterative prosodic structure than with a recursive one. This is called here "Assumption 2".

Ladd (1986) and Hunyadi (2006) advance models incorporating the prediction of Assumption 1 and motivating recursion. By contrast, a model corresponding to Assumption 2 cannot be assumed to be recursive.

In this paper, we do not study parenthetical clauses. Rather we concentrate on relative clauses. Although the relative clauses used here are ambiguous between a restrictive and a non-restrictive reading, they are semantically closer to the head noun than parenthetical clauses, which are independent comments

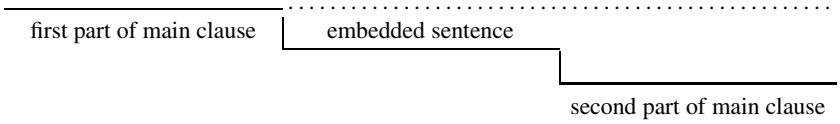

Figure 2: Assumption 2: Regular downstep across the whole sentence (iterative prosody)

(Lehmann 1986). In other words, we expect a tighter prosodic cohesion between relative clause and matrix clause than between clausal parenthetical and matrix clause, like those studied by Kutik et al. and by Ladd. The relationship between scaling of tones in sentences with embedded relative clauses can provide valuable insights as to the recursive potentials of higher prosodic domains, because, in our view, they interrupt a larger domain of the same kind. Since the domains we are interested in are all clauses, they all form domains at the same level of prosodic structure, which we call here p-phrases (see Selkirk 2009 for Match constraints). If the interruption in the prosodic structure of a matrix clause has no other effect than an additional downstep, comparable to the pattern found in lists (see (1)), we should conclude that no recursion is present in the prosodic structure. However, if the second part of the clause is scaled relatively to the first one, independently of the interruption, and in spite of the fact that the two parts are separated by embedded material, then the existence of recursion in prosodic structure should be considered confirmed.

1.3. Typological comparison

German and Hindi are compared as to their potential prosodic recursivity at the level of p-phrases. German is a so-called intonation language (see Ladd 2008, Gussenhoven 2004, Jun 2005 for typologies of intonation systems), in which pitch accents and boundary tones are realized as a function of syntax, information structure, scope relationship and other pragmatic needs. Prosodic phrasing is erected from syntactic structure, and the scaling of tonal events is primarily guided by this syntax-driven phrasing. In such a language, pitch values are scaled to the pitch register of the prosodic domain, which itself is relative to the other prosodic domains of the sentence. A focused word raises pitch, whereas givenness lowers it. The properties of the tonal structure of German are well understood and are similar to those of English.

Some intonational properties of Hindi are also well studied (see for instance Harnsberger and Judge 1996, Patil et al. 2008). Hindi bases its intonation on prosodic domains, the delimitation of which is also driven by syntax. Each domain is characterized by an initial rising tonal event, LH, and a final rising

event, LH, aligned with the boundaries of the domains rather than with lexically accented syllables, as in German. The two rises are often realized only partially, and in some cases they are even absent, depending on style and tempo among other factors. The intonation of Hindi is thus based on its prosodic phrasing to a much greater extent than the intonation of German is, since all tonal events are related to prosodic phrasing. Hindi is a typical 'phrase language' (see Féry 2010). Pitch scaling is also an important aspect of Hindi intonation, post-focal compression being the most conspicuous effect in this respect (see Patil et al. 2008). Pre-focal or focal domains do not change their height. Downstep is not canceled as easily as in German. But altogether, many aspects of Hindi intonation remain to be studied, and basic issues, such as the existence of lexical stress, are not fully understood.

This paper compares the results for the intonational correlates of the same kind of syntactically embedded material in these two languages. In the following sections, each experiment is described and the results are discussed, first for German and then for Hindi (Sections 2 through 5). In Section 6, a comparison is made and the implications for the theoretical modeling of prosody are discussed. Section 7 concludes the paper.

2. German experiment

2.1. Stimuli

In a production experiment, the prosody of sentences with center-embedded clauses was investigated. The same items were recorded without embeddings for comparison. As exemplified in (7a), the recursive sentences involved copulative constructions with a center-embedded relative clause after the subject noun ([$_{CP1}$]), in which a further relative clause was center-embedded ([$_{CP2}$]). The counterpart without embeddings appears in (7b).

(7) a. [*Die Birnen*$_{Subj}$ [$_{CP1}$ *die an dem Baum* [$_{CP2}$ *der*
the pears which at the tree which
grün ist] *hängen*] *sind sauer* $_{VP}$]
green is hang are sour
'The pears that hang from the tree that is green are sour.'

b. *Die Birnen sind sauer.*
the pears are sour
'The pears are sour.'

The set of test items comprised a total of twelve sentences: six recursive sentences and their counterparts without embedding. The complete set of experimental sentences appears in Appendix A.

2.2. Recordings

The production experiment involved the audio recordings of 21 female native speakers of German aged between twenty and thirty, all students at the University of Potsdam. The recordings took place individually in a soundproof booth. The sentences were pseudo-randomized and presented to the participants in a Powerpoint presentation (one item per slide), in which the target items were separated by one or two fillers. The sentences were recorded on a DAT recorder. The participants were instructed to read the sentences at a normal speech rate and as naturally as possible. If the participants felt that they had failed to produce an utterance correctly, they were given the chance to correct themselves. In this case, only the corrected sentence was included in the analysis. A total of 252 utterances were recorded and analyzed.

2.3. Measurements

The analysis of the data was carried out with the acoustic analysis software Praat (Boersma and Weenink 1993–2010). The recorded utterances were subdivided manually into subsequences. The subject noun and the first word of the matrix VP were the target words for measurements (underlined in (8a, b)). In order to render the sentences comparable, the target words were kept constant as far as the number of syllables and stress are concerned: the subject nouns were disyllabic with stress on the initial syllable and a schwa in the second syllable. All their segments were voiced. In all sentences with relative clauses, the first word of the matrix VP was the main verb *sind* 'are'.

A boundary was set at each point of embedding and 'de-embedding'[2] (in the terminology of Hunyadi 2006) as shown in (8a). The respective counterparts without the insertion (8b) were subdivided into subject NP and VP.

(8) a. Die Birnen | die an dem Baum | der grün ist | hängen | sind sauer
L*H$_P$ L*H$_P$ L*H$_P$ L*H$_P$ H*L$_I$

b. Die Birnen | sind sauer
L*H$_P$ H*L$_I$

In all sentences, the peak values of the subject noun and of the first word of the matrix VP were measured with the help of a Praat script. These values corresponded to the H of the bitonal L*H$_P$ accent realized on the subject NP, and to the beginning of the matrix VP.[3] The dots in Figure 3 mark the points

2. De-embedding is the position at which an embedded clause ends.
3. If no voicing was present at a particular boundary, the point of voicing onset in the target word was measured.

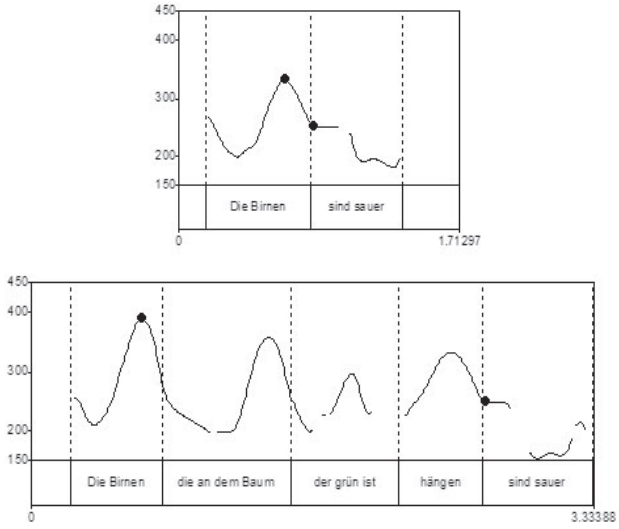

Figure 3: Exemplary pitch tracks of sentences without (8a) and with (8b) embedded relative clauses

of measurement. In nearly all utterances, the accented word in the VP – the adjective *sauer* in (8a, b) – was lower than the auxiliary at the beginning of the VP. In a few sentences, the adjective was higher, but we decided to measure the same position in all cases because this value is the one which establishes a relation between the two parts of the matrix clause.

Furthermore, we measured the peak values of the embedded material. The first part of the first embedding (CP1), the second embedding (CP2) and the second part of the first embedding (CP1) each had one L*H accent, the peaks of which were the points of measurement. These accents were consistently realized on the noun in the first part of embedding 1, on the adjective in embedding 2 and on the verb in the second part of embedding 1 (see Figure 3). We had to manually correct the outputs of the script in approximately 40 % of the cases. In cases of microprosodic distortions, the F_0 values were measured at a place nearby.

For each type of item, the values of all sentences were averaged across all speakers.

3. Results for German

3.1. Results of F_0 measurements in the main clause

Figure 4 shows the difference between the F_0 values of the main clauses in sentences with two embedded sentences and those without embedded sentences. The averaged results for the measurements are plotted in a graph. The difference in pitch between the highest tone of the subject and the beginning of the VP differs considerably in the two types of sentences. In the sentences with embedded clauses, the average difference is 67 Hz while it is only 18 Hz in the sentences without embedding. Thus on average there is a difference of 49 Hz. It should be noted that a fair amount of variation is apparent among the sentences, although the difference between the two measurement points is always larger in the sentences with embedding than in the sentences without embedding. Furthermore, not only are the high values on the subject NP higher when embeddings are added, but the low values at the beginning of the VP are also lower.[4]

Table 1 sums up the results for the main clauses in all sentences.

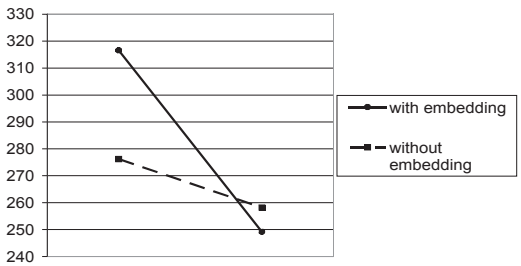

Figure 4: Measured tonal values in all four types of sentences

Table 1: Overall average pitch values of the main clauses

	With embedding	Without embedding
Subject	316 Hz	276 Hz
VP	249 Hz	258 Hz

4. The paired t-test for the low values with and without embedding is highly significant: $p = 0.0003$.

3.2. Pitch results for the embedded sentences

Figure 5 displays the average measurement points in the sentences with center-embedded relative clauses in the form of a diagram. It shows that the pitch register of the embedded material is much narrower in comparison with the overall pitch register of the sentences. In the first part of the main clause, the average highest point is 316 Hz, and in the second part of the main clause, the last measured point in the sentence is 249 Hz. The total range of the main clause is thus 67 Hz. In the embedded clauses, the average highest point ranges from 274 Hz in the first part of embedding 1 to 257 Hz in embedding 2, thus a difference of 17 Hz between the first and the second embedded clause. The value of the second part of embedding 1 (267 Hz) is higher than that of embedding 2, but lower than that of the first part of embedding 1.

Summing up, we observe a large downstep between the main clause and embedding 1, a smaller downstep between embedding 1 and embedding 2 and a reset between embedding 2 and the second part of embedding 1. At the end of the sentence, downstep is again observed.[5] The most interesting result is the position of measurement 4, which is the second measurement of the first relative clause, following the most deeply embedded relative clause.[6]

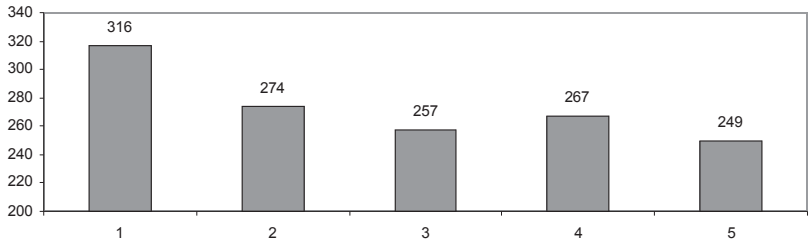

Figure 5: Diagram of the average Hz values of the measurement points in sentences with embeddings

5. Given the upstep between the second embedded clause and the end of the first embedded clause, the same upstep effect could be expected between the end of the first embedded clause and the end of the main clause. But this is not what we find, since there, a regular downstep takes place. We assume that this is due to the fact that the end of the sentence has been reached. This part of the sentence is short, and has a falling contour. This is a place where neutralization of tone value takes place.
6. This point of measurement is just an average and cannot be seen as a value realized by all speakers. Some speakers do realize a downstep in all parts of the sentence, and others have an upstep at point 4 which is higher than the point of measurement 2. However, taken together, the average pattern is compatible with prosodic recursion. Clearly more studies are needed to elucidate the role of pitch scaling as an indicator of syntactic recursion.

3.3. Discussion of the German results

The results obtained for the tonal scaling of the two parts of the main clause do not confirm Assumption 1, namely that the embedded part of the sentence is invisible for the matrix clause, at least not for the speakers who realized an upstep on the second part of the first embedded clause. Recall that this hypothesis posited that there is no difference between the tonal scaling of the matrix clause when it is continuous and when it is separated by embedded material. This was illustrated in Figure 1. Indeed, in the German sentences, a clear effect of length was found. The longer sentences with embedded clauses started higher and ended lower than the same sentences without embeddings. In the experimental sentences without embedding, the difference in the tonal height between the subject and the verb exhibited an average of 18 Hz, but in the sentences with embedding, the tonal lapse was 67 Hz, a large difference (see Table 1). Perceptively, the effect of cutting out the embedded material renders the remaining sentences bizarre, as if the speaker starts with a narrow focus on the subject and changes her mind in the VP part.

Assumption 2, which posited that the second part of a clause is scaled relatively to the preceding clause (see Figure 2), is not wrong, but it is insufficient as an explanation of our data. According to Assumption 2, a difference between the two versions of the sentences – with and without embedding – was expected in the scaling of the second part of the sentence as compared to the embedded part. The first part of the main clause was assumed to be identical in the two kinds of sentences. This expectation was not fulfilled. As is clearly visible from Figure 5, both the value of the subject and that of the verb differ from each other. In the sentences with embedded material, the F_0 value of the subject is higher than in sentences without embedding, and the second measured point is lower. In other words, the entire range used in the sentences with embedding is larger than that used in the sentences without embedding.

The second difference between the predictions of Assumption 2 and our results has to do with tone scaling, i.e., with downstep and upstep in the different parts of the sentences containing embeddings. The second part of the first relative clause was upstepped relatively to the second relative clause but downstepped relatively to the first part of the first relative clause. This is not predicted by Assumption 2, and speaks for a more complex relationship between syntax and prosody than Assumptions 1 and 2 are able to predict. We will return to this issue in Section 6. But first, let us turn to Hindi.

4. Hindi experiment

Sentences with two center-embedded clauses were recorded in Hindi as part of a larger experiment including different kinds of production sentences. No simple sentences (without center-embedded clauses) were recorded. The experiment was carried out as a Powerpoint presentation. As in the German experiment, participants were instructed to produce the sentences displayed on the screen as naturally as possible. Once they answered the question without any hesitation the next stimulus was presented. If the speakers showed any hesitation, they were asked to repeat the sentence. The presentation flow was controlled by the experimenter, but the participants were allowed to take a few minutes break when they wanted to do so. A set of instructions and two practice examples familiarized the subjects with the process.

The entire data set used in this experiment was collected in one experiment, run individually with 30 speakers. All were female students, and they were recorded at the University of Delhi on a DAT recorder. They were native speakers of Hindi in their twenties, coming mostly from Delhi and the surrounding states. They were paid for their participation. Out of 30 speakers, data from only 20 speakers was used for the analysis. The selection of the speakers was not based on any criteria, but was completely random. A total of 120 sentences (20 speakers $\times$ 6 items) were analyzed. The complete set of sentences appears in Appendix B.

An example of a sentence is given in (9).

(9) Example of a Hindi sentence
[*Woh naukar*]$_1$ [*jisko us phalavaale ne*]$_2$ [*jo*
that servant whom that fruit-seller ERG who
bhaag rahaa thaa]$_3$ [*jagaayaa*]$_2$ [*gariib thaa*]$_1$
running was woke-up poor was
'The servant whom the fruit seller who was running woke up was poor.'

The measurements performed on the Hindi sentences are illustrated in Figure 6, an annotated pitch track. They were more numerous than in the German ones, because of our lack of experience with Hindi prosody, and our consequent concern about missing important points of measurements. As with the German data, the recorded utterances were subdivided manually into subsequences, marked with numbers in Figure 6. In every numbered domain, two measurements were performed: the first F_0 minimum and the following F_0 maximum. In many cases, one of the low tones was missing. This happened when the portion considered ended in a high tone, a pattern which was pervasive due to the tonal structure of Hindi (see Patil et al. 2008). Analysis of the data was again carried out with the acoustic analysis software Praat (Boersma

and Weenink 1993–2010) and corrected manually when necessary. As with the German data, the values measured for each position of the domains were averaged for all speakers and sentences.

5. Results for Hindi

5.1. Results of F_0 measurements

The results of the F_0 measurements are displayed in Figure 7. Each peak corresponds to the highest tone in a domain, as indicated in Figure 6, and each low point in-between is the lowest point between the high points of two domains.

As is visible from the figures, the pattern differs from that of German. There is a downstep in every clause, which takes place inside of the clause. However, the first measured high point in every clause is more or less as high as the last measured high point of the preceding one. No reset occurs at the beginning of the second part of the relative clause (the second CP1), even though the value at this point is as high as the last value of CP2. Compared to the first value of CP2, this point is clearly downstepped. The same is true of the transition between relative clause 1 and relative clause 2. Here, too, the two adjacent values are more or less equal. But compared to the first value of CP1, the relative clause CP2 is downstepped. Inside of clause, the two high measurement points show downstep. And of course, if one only compares the first H of the embedded clause, or the last H of every embedded clause, we also find downstep. However there is no evidence of reset.

5.2. Discussion of the Hindi results

For Hindi, Assumption 2 makes the right predictions. In this language, there is no reason to assume a recursive prosody, at least in the data analyzed in this paper. The sentences are realized with a regular downstep on all parts. As a result, a simple iteration of prosodic phrases seems to be sufficient to account for the pattern.

6. Discussion

The results of the experiments in German and Hindi have clear implications for the interface between syntax and prosody. The main issue that is addressed in this paper, namely the question of the recursion of prosodic domains, could not be answered uniformly for both languages. This section discusses some of

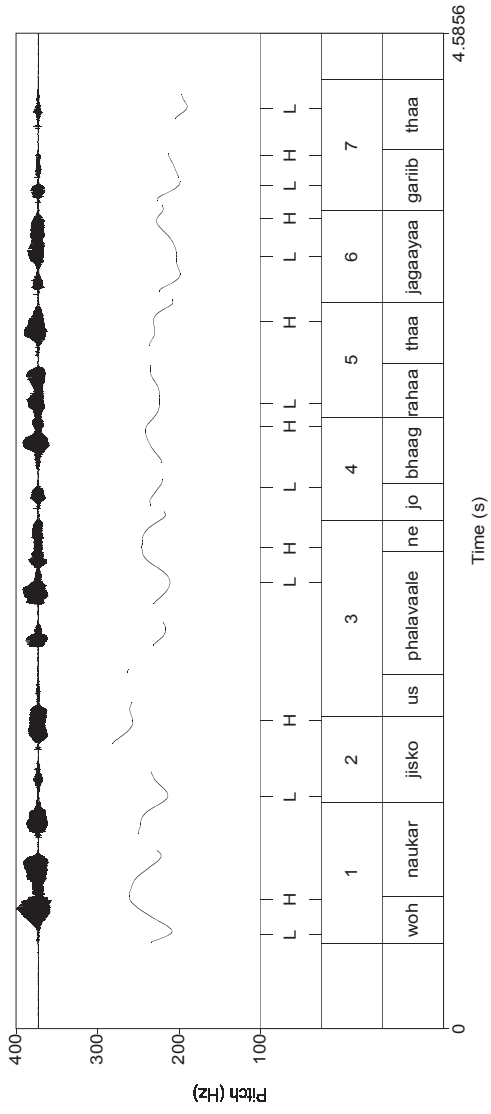

Figure 6: Pitch track of a Hindi sentence

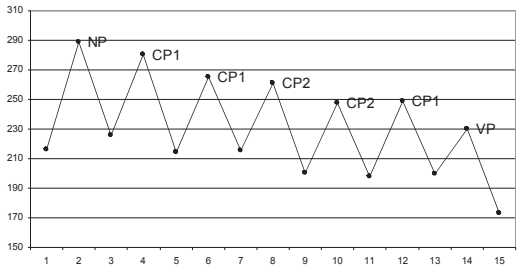

Figure 7: High and low tones of the Hindi sentences

the implications of the obtained effects for the theoretical modeling of prosody (especially pitch scaling) in the two languages.

6.1. Iterativity of prosodic phrasing: A model for Hindi

We could not confirm for German that the effect of removing a relative clause from a main clause had no influence on the prosody of its two parts (contrary to Hunyadi 2006, who finds this for Hungarian relative clauses, and to Kutik et al. 1983, who describe this for English parenthetical clauses). Thus Assumption 1, illustrated in Figure 1, could not be confirmed.

Assumption 2 in Figure 2 shows an alternative model for pitch scaling in which a non-initial prosodic phrase is always downstepped in the same way relatively to a preceding one, regardless of the syntactic structure. This model is compatible with a view of iterative prosodic phrasing: every prosodic phrase stands for itself, and is scaled relatively to the preceding one. Reference lines of prosodic domains at a certain distance are not available for tone scaling.

This view of prosody makes the right predictions for Hindi, where every prosodic domain is scaled relatively to the immediately preceding one. There was no indication that the first part of a clause which was separated from its second part by embedded material was still available for pitch adjustment. Sentence (9) is reproduced in (10) with the suggested prosodic structure.

(10) Prosodic phrasing of Hindi
[*Woh naukar*]$_P$ [*jisko us phalavaale ne*]$_P$ [*jo*
that servant whom that fruit-seller ERG who
bhaag rahaa thaa]$_P$ [*jagaayaa*]$_P$ [*gariib thaa*]$_P$
running was woke-up poor was
'The servant whom the fruit seller who was running woke up was poor.'

The result for Hindi is preliminary and needs to be investigated and confirmed in further experiments. In particular, it can be seen in Figure 7 that downstep is larger within a clause than across clauses, which speaks for a different sensitivity of prosody to syntax than that found in German.

6.2. Recursive prosody: A model for German

The results for German tell another story, since there, the tonal relationship between the different parts of the sentence was more complex. At the beginning of the sentence, the matrix clause, the first part of the first relative clause and the first part of the second relative clause were in a downstep relationship to each other. However, the second part of the first relative clause was upstepped relatively to the second relative clause (in the average values), but downstepped relatively to the first part of the first relative clause. We can thus assume that the reference line of the first part of a clause still counted for the second part, at least for those speakers who realized an upstep; see Figure 8 below. In German, neither Assumption 1 nor Assumption 2 are perfect predictors of the results. Assumption 1 predicts that the embedded relative clauses are invisible for the matrix clause, a wrong result for German. And Assumption 2 says that downstep happens iteratively regardless of the embedding of syntactic structure. Again, this is not what was found.

Because of the difference between the two languages, a unique prosodic structure for all cases does not seem possible. We suspect that the upstep pattern at the de-embedding location in German as compared to the absence of upstep in Hindi reflects a crucial difference in the prosodic organization of the two languages.

Two elements in our experimental results have to be taken into account in a model of German prosody: first, the fact that the sentence-initial high tones are raised when the sentences have embedded relative clauses as compared to their counterparts without embedding; and second, the fact that the scaling of the second part of the first relative clause which interrupts downstep is adjusted to the reference line of the first part of CP1.

The first property can be explained by what we call register enlarging, formulated in (11).

(11) Register enlarging
If a non-recursive p-phrase has a range called r, the same phrase containing embedded p-phrases is enlarged by a factor f, and embedded p-phrases have a narrower register, decreased relatively to r by a factor $-f$.

Register enlarging is expressed in terms of prosodic phrases, and claims that a prosodic phrase which contains another prosodic phrase of the same level

has an enlarged register as compared to its equivalent without embedding. This accounts for the values in Table 1. The idea is that if a main clause contains relative clauses, then its register is expanded in order to be able to contain the entire register of the embedded clauses. And as a result of embedding, the relative clauses themselves have a reduced register.

We cannot exclude the possibility that at least part of the enlarging effect correlates with length. In this sense, a longer sentence starts higher than a shorter one.[7] The final lowest point in a sentence is not as variable as the initial highest point, and neutralization of the low target of a declarative sentence seems to be the rule (see also Féry and Kentner 2010 for neutralization of the final low target). However, the data reported here also reveal that embedded clauses do not just decrease their register in a linear way but are sensitive to the prosodic embedding.

To explain the upstepped value in the second part of CP1, we adopt a model of pitch scaling in which prosodic domains at the same level n are in a downstep relation to each other (see also van de Berg et al. 1992, Truckenbrodt 2007, Féry and Truckenbrodt 2005, Féry and Ishihara 2009).[8] We focus here on the top reference lines to which high tones are scaled. At lower levels, the prosodic phrases are also in a downstep relationship to each other, leading in certain cases to an intricate pattern of downstep. Figure 8 illustrates the downstep pattern in our sentences, with reference lines available until the end of the sentence. The main aspect of this pattern leading to the assumption of recursive prosodic structure is that the second part of CP1 is adjusted to the reference line of the first relative clause, causing an upstep from the level of CP2, and disrupting in this way the regular downstep of the sentence. Figure 8 also shows the lowest bottom line, to which the final matrix VP is scaled.

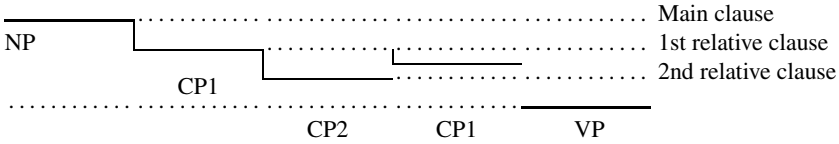

Figure 8: Recursive p-phrasing in German

7. In order to eliminate the hypothesis that only length is responsible for the upstep effect found between the two last parts of the relative clause, additional studies comparing sentences with embedded clauses and root sentences without embedding will be conducted in future research.
8. The behavior of the final part of the main clause cannot be accommodated easily in this model, as it seems to be scaled relatively to the second part of the first embedded clause, rather than to the first part of the main clause. We suspect a final lowering effect, but more research is needed to clarify this point.

We propose that embedding of prosodic domains, as illustrated in (4) and repeated in (12), best accounts for the data obtained for German. In this kind of prosodic phrasing, small p-phrases are contained in larger ones, and the prosodic structure as a whole is recursive.

(12) [*Die Jungen* [*die das Kind,* [*das Angst hat*]$_P$, *ärgern*]$_P$
the boys who the child who fear has tease
sind böse]$_P$
are nasty
'The boys who tease the child who is scared are nasty.'

By contrast, the prosodic structure in Hindi is iterative, as illustrated in (10), repeated in (13).

(13) [*Woh naukar*]$_P$ [*jisko us phalavaale ne*]$_P$ [*jo*
that servant whom that fruit-seller ERG who
bhaag rahaa thaa]$_P$ [*jagaayaa*]$_P$ [*gariib thaa*]$_P$
running was woke-up poor was
'The servant whom the fruit seller who was running woke up was poor.'

In short, the prosodic structures that we propose for German and Hindi differ in one main respect: German allows recursion and embedding of p-structure and mimics in this way the syntactic structure, whereas Hindi does not allow recursion, but simply iterates its prosodic domain.

6.3. Implications for the Strict Layer Hypothesis (SLH)

In Section 1.2, the SLH was addressed, which only allows a constituent p to be exhaustively dominated by a constituent of the immediately higher level p+n. Some aspects of the SLH remain uncontroversial, as for instance what can be called the unidirectional orientation of the prosodic hierarchy. Constituents cannot be dominated by constituents of a lower level than themselves (p cannot be dominated by p−n). Other assumptions of the SLH were claimed to be in need of revision on principled grounds. We showed in section 1.2 that the prosodic hierarchy does not have enough categories to be maintained in complex sentences.

In the present paper, we have shown that the tonal scaling of sentences with embedded relative clauses brings a strong argument in favor of recursive prosodic structure. A strict application of the SLH would imply that all parts of sentences (12) and (13) are phrased independently at the level of p-phrases (or intermediate phrases), which in turn would imply that Assumption 2 is correct. This is what was found for Hindi, but not for German.

Let us assume that downstep applies between the constituents of a phrase. As a result, downstep is an indicator of prosodic structure. We can now test several hypotheses. If we schematize the structure in (12) and (13) with letters (A corresponds to the first part of the matrix clause, B to the first part of the first relative clause, C to the first part of the second relative clause, etc.), several options are available as to how the constituents can be organized prosodically, as shown in (14).[9] In (14a), all constituents are parts of the same level of phrasing, a p-phrase or an i-phrase or both. Such a structure is expected to display downstep on all its parts (Liberman and Pierrehumbert 1984 and many others after them). A structure like the one in (14b) does not differ much from (14a) in terms of its downstep predictions. Here, every constituent is phrased individually, but since they are all part of a unique intonation phrase, and since iteration is the unique structure, downstep is also expected.

In (14c), by contrast, additional prosodic structure changes the expectations as far as downstep is concerned. (14d) is a simplification of (14c), in that all individual phrases indicated in (14c) are not visible. It is thus much easier to read. If we go from the most deeply embedded constituents outwards, C and D are both phrased individually (as are all other parts of the sentence), but they are also phrased together in the constituent P3. Downstep is thus expected between C and D (correctly found in Figure 5 for German). Going one level up, B and E are phrased together in P2, and thus a downstep is expected between B and E, which is also confirmed in Figure 5. But now, B and E are no longer adjacent, since P3 is intervening. For this reason, downstep is expected between B and C, the beginning of constituent P3. Again this is borne out. No downstep is expected between D and E, because D is inside of constituent P3, and invisible to E. This is again exactly what we found in Figure 5. The outer layer P1 behaves in the same way as far as the initial part is concerned: a downstep is realized between A and B. However, since F is subject to final neutralization, the expected upstep is not found.

(14) a. $[\ [A\ B\ C\ D\ E\ F]_{P}\ ]_{I}$
b. $[\ [A]_{P}\ [B]_{P}\ [C]_{P}\ [D]_{P}\ [E]_{P}\ [F]_{P}\ ]_{I}$
c. $[\ [_{P1}\ [A]_{P}\ [_{P2}\ [B]_{P}\ [_{P3}\ [C]_{P}\ [D]_{P}\ ]_{P3}\ [E]_{P}\ ]_{P2}\ [F]_{P}\ ]_{P1}\ ]_{I}$
d. $[\ [_{P1}\ A\ [_{P2}\ B\ [_{P3}\ C\ D\ ]_{P3}\ E\ ]_{P2}\ F\ ]_{P1}\ ]_{I}$

To sum up, the structure in (14c) is an input to tonal scaling, the larger prosodic domains being influenced by their parts and by their neighbors. And the syntactic structure is an input to the prosodic structure. It is the recursive syntactic

9. We are grateful to an anonymous reviewer who proposed this notation as a basis for the discussion of the implications of the SLH.

structure which is the trigger for the recursive prosodic structure (see Schubö 2010 for this conclusion). And the recursive prosodic structure is an input to the phonetic tonal scaling.

We agree with Selkirk's (2009) Match constraints, which allow recursion of prosodic domains by default, as well as with Ito & Mester's (2008, to appear) minimal and maximal p-phrases. However, we are unable to show how these theoretical notions are implemented in our data due to space limitations.

7. Conclusion

Sentences containing center-embedded relative clauses were explored in a production experiment in German and in Hindi. For German, the complex sentences were compared with the same sentences without embeddings. The aim of this study was to understand whether syntactic recursion is mapped to a similarly recursive prosodic structure, and whether this is a universal property. The results revealed a significant difference between the two languages.

In German, the values of the initial pitch accents were lower in sentences without embedding than in sentences with embedding, whereas the final pitch accents were higher, showing that the overall register used was reduced in the short sentences. Contrary to what has been described for Hungarian (Hunyadi 2006), it was not the case that the second part of a main clause started at the same level as in a sentence containing embedded material. On the contrary, the effect of removing the embedded material from a long sentence was that the sentence sounded awkward, as if the subject was pronounced much too high and/or the VP much too low.

The main result of the experiment, however, lies in the interesting downstep and upstep pattern in the sentences with embedding. In German, the second part of the first relative clause was reset relatively to the second relative clause, but downstepped relatively to the first part of the first relative clause. We interpret this result as an indication that the prosodic structure in German is recursive, in the sense that one prosodic phrase of a certain level n can be interrupted by an embedded prosodic phrase of the same level n.

In Hindi, we found a different pattern. Here, it was not the case that the second part of the first relative clause was reset after the second relative clause. All parts of the sentence were in a downstep relationship to each other (considering only the initial high tones of each part). We interpreted this pattern as non-recursive. The prosodic structure of Hindi is in line with an iterative view of prosodic structure, in which the prosodic domains follow each other but do not overlap.

The difference between the two languages confirms previous experiments on Hindi (Patil et al. 2008) showing that the prosodic structure of this language

truly differs from results for German. It correlates with the assumption that Hindi intonation is 'phrase-based,' rather than pitch-accent based as in German.

University of Frankfurt/Main
caroline.fery@googlemail.com
University of Potsdam
fabian.schuboe@googlemail.com

Appendix

A. German test items

(1) a. *Die Birnen, die an dem Baum, der grün ist, hängen, sind sauer.*
the pears that at the tree that green is hang are sour
'The pears that hang from the tree that is green are sour.'

b. *Die Birnen sind sauer.*
the pears are sour
'The pears are sour.'

(2) a. *Die Hosen, die der Mann, der arm ist, trägt, sind billig.*
the trousers that the man who poor is wears are cheap
'The trousers that the man who is poor wears are cheap.'

b. *Die Hosen sind billig.*
the trousers are cheap
'The trousers are cheap.'

(3) a. *Die Jungen, die das Kind, das Angst hat, ärgern, sind böse.*
the boys who the child who fear has tease are nasty
'The boys who tease the child who is scared are nasty.'

b. *Die Jungen sind böse.*
the boys are nasty
'The boys are nasty.'

(4) a. *Die Bären, die im Wald, der naturbelassen ist,*
the bears that in the forest which nature-left is

leben, sind friedlich.
live are peaceful
'The bears that live in the forest which has been left in a natural state are peaceful.'

b. *Die Bären sind friedlich.*
the bears are peaceful
'The bears are peaceful.'

(5) a. *Die Blumen, die der Junge, der einsam ist, sammelt, sind rosa.*
the flowers that the boy who lonely is collects are pink
'The flowers that the boy who is lonely collects are pink.'

b. *Die Blumen sind rosa.*
the flowers are pink
'The flowers are pink.'

(6) a. *Die Bauern, die den Mais, der reif ist, ernten, sind emsig.*
the farmers who the corn that ripe is harvest are industrious
'The farmers who harvest the corn that is ripe are industrious.'

b. *Die Bauern sind emsig.*
the farmers are industrious
'The farmers are industrious.'

B. Hindi test items

(1) *Woh naukar jisko us phalavaale ne jo bhaag rahaa thaa jagaayaa gariib thaa.*
that servant whom that fruit-seller ERG who running was woke-up poor was
'The servant whom the fruit seller who was running woke up was poor.'

(2) *Woh baalak jisko us aadamii ne jo gaa rahaa thaa bulaayaa dubalaa thaa.*
that child whom that man ERG who singing was called weak was
'The child whom the man who was singing called was weak.'

(3) *Woh mainejar jisko us pulisawaale ne jo chal rahaa thaa maaraa amiir thaa.*
that manager whom that policeman ERG who walking was beat rich was
'The manager whom the policeman who was walking beat was rich.'

(4) *Woh muniim jisko us graahak ne jo bol rahaa thaa*
that clerk whom that customer ERG who talking was
pukaaraa motaa thaa.
called fat was
'The clerk whom the customer who was talking called was fat.'

(5) *Woh dukaanadaar jisko us kisaan ne jo ro rahaa*
that shopkeeper whom that farmer ERG who crying
thaa manaayaa badhiir thaa.
was persuaded deaf was
'The shopkeeper whom the farmer who was crying persuaded was deaf.'

(6) *Woh gaayak jisko us sipaahii ne jo soch rahaa thaa*
that singer whom that soldier ERG who thinking was
giraayaa biimaar thaa.
pushed ill was
'The singer whom the soldier who was thinking pushed was ill.'

References

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*, 335–367. Cambridge: Cambridge University Press.

Boersma, Paul & David Weenink. 1993–2010. Praat: Doing phonetics by computer [Computer program].

Cooper, William & John Sorensen. 1981. *Fundamental frequency in sentence production*. Heidelberg: Springer.

Féry, Caroline. 1993. *German intonational patterns* (Linguistische Arbeiten 285). Tübingen: Niemeyer.

Féry, Caroline. 2010. The intonation of Indian languages: An areal phenomenon. In Imtiaz Hasnain & Shreesh Chaudhury (eds.), *Problematizing Language Studies. A Festschrift in honour of Ramakant Agnihotri*, 288–312. New Delhi: Aakar Books.

Féry, Caroline & Shinichiro Ishihara. 2009. How Focus and Givenness shape prosody. In Malte Zimmermann & Caroline Féry (eds.), *Information structure from different perspectives*, 36–63. Oxford: Oxford University Press.

Féry, Caroline & Gerrit Kentner. 2010. The prosody of grouping names in German and Hindi. *Proceedings of Speech Prosody*, Chicago. 1–4.

Féry, Caroline & Hubert Truckenbrodt. 2005. Sisterhood and tonal scaling. In Merle Horne & Marc Oostendorp (eds.), *Boundaries in intonational phonology* (Studia Linguistica 59), 223–243.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Harnsberger, James D. & Jasmeet Judge. 1996. Pitch range and focus in Hindi. *The Journal of the Acoustical Society of America* 99. 2493.

Hauser, Marc D., Noam Chomsky, W. Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298. 2769–2779.

Hunyadi, László. 2006. Grouping, the cognitive basis for recursion in language. *Argumentum* 2. 67–114.

Ito, Junko & Armin Mester. 2008. The extended prosodic word. In Janet Grijzenhout & Baris Kabak (eds.), *Phonological domains: Universals and deviations*, 105–194. Berlin: Mouton de Gruyter.

Ito, Junko & Armin Mester. to appear. Recursive prosodic phrasing in Japanese. In Toni Borowsky, Shigeto Kawahara, Takahito Shinya & Mariko Sugahara (eds.), *Prosody matters: Essays in honor of Elisabeth Selkirk*. London: Equinox Publishers.

Jun, Sun-Ah (ed.). 2005. *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press.

Kutik, Elanah J., William Cooper & Suzanne Boyce. 1983. Declination of fundamental frequency in speakers' production of parenthetical and main clauses. *Journal of Acoustics Society of America* 73. 1731–1738.

Ladd, D. Robert. 1986. Intonational phrasing: The case of recursive prosodic structure. *Phonology Yearbook* 3. 311–340.

Ladd, D. Robert. 1990. Metrical representation of pitch register. In John Kingston, & Mary E. Beckman (eds.), *Papers in laboratory phonology 1: Between the grammar and the physics of speech*, 35–57. Cambridge: Cambridge University Press.

Ladd, D. Robert. 1992. *Compound prosodic domains*. Occasional Papers of the Linguistics Department. University of Edinburgh.

Ladd, D. Robert. 2008. *Intonational phonology*. 2nd edn. (1st edn. 1996). Cambridge: Cambridge University Press

Lehmann, Christian. 1986. On the typology of relative clauses. *Linguistics* 24. 663–680.

Liberman, Mark & Janet Pierrehumbert. 1984. Intonational invariance under changes in pitch range and length. In Mark Aronoff & Richard Oehrle (eds.), *Language sound structure*, 157–233. Cambridge: MIT Press,

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

Patil, Umesh, Anja Gollard, Gerrit Kentner, Frank Kügler, Caroline Féry & Shravan Vasishth. 2007. Focus, word order and intonation in Hindi. *Journal of South Asian Linguistics* 1. 53–70.

Pierrehumbert, Janet & Mary Beckman. 1988. *Japanese tone structure* (Linguistic Inquiry Monograph 15). Cambridge: MIT Press.

Schubö, Fabian. 2010. *Recursion and prosodic structure in German complex sentences*. Potsdam: University of Potsdam MA thesis.

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. 35–73.

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

Wagner, Michael. 2005. *Prosody and recursion*. Cambridge, MA: MIT dissertation.

Watson, Duane & Edward Gibson. 2004. The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes* 19(6). 713–755.