SISTERHOOD AND TONAL SCALING*

Caroline Féry & Hubert Truckenbrodt

Abstract. This paper discusses central aspects of the effects of hierarchical structure on tonal scaling in intonation. The core results of a number of phonetic studies on this topic, by Ladd, by van den Berg, Gussenhoven and Rietveld, as well as experimental results of our own, are reviewed. We review the suggestions of this earlier work and argue for an addition to the theory. The principle 'The deeper the steeper' says that downstep among sister nodes is relatively larger if these sister-nodes are relatively more deeply embedded in the prosodic representation.

1. Introduction

In our earlier works on German intonation, we pursued topics like the role of prosodic categories and information structure (Féry 1992, 1993), aspects of the relation of syntactic structure to intonation (Féry & Hartmann 2005), as well as effects of prosodic categories on aspects of tonal scaling (Truckenbrodt 2002, to appear).

In our present joint work, we are investigating issues in the related topic of the connection between hierarchical structure and tonal scaling. The experiment reported in Truckenbrodt and Féry (2004) sought to replicate, with German data, the results of the important study by Ladd (1988). The present paper discusses some theoretical issues related to this experiment and the work on which it was based. Some of the main results of Ladd (1988), the likewise relevant results of van den Berg et al. (1992) and Truckenbrodt (2002), as well as of our experimental study are reported, and some ways of putting together these proposals in connection with hierarchical structure and tonal scaling are suggested.

2. Background

2.1. *Downstep, partial reset and embedded register levels: Ladd's (1988) experiments*

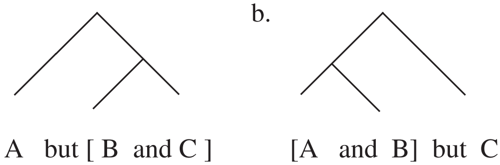
Ladd (1988) compared complex sentences with the two different structures in (1), where A, B, and C are clauses of similar rhythmic and syntactic pattern, with clause-internal downstep. An example of each

*We would like to thank the audience of the workshop on intonational boundaries for helpful feedback, and to an anonymous reviewer for helpful comments. The first author's contribution to the work presented here is part of Project A1 of the Sonderforschungsbereich on Information structure at the University of Potsdam (SFB 632). The joint project has been conducted while the second author was a guest professor at the Center of Cognitive Studies in Potsdam during the winter semester 2002–2003.

Studia Linguistica 59(2/3) 2005, pp. 223–243. © The Editorial Board of Studia Linguistica 2005.
Published by Blackwell Publishing Ltd., 9600 Garsington Road, Oxford OX4 2DQ, UK, and
350 Main Street, Malden, MA 02148, USA

case is given in (2). The clauses contained one of three subjects (*Ryan*, *Warren*, and *Allen*) and one of three predicates (*has a lot more money*, *has more popular policies* and *is a stronger campaigner*). In a second experiment, the subjects were changed to *Governor Allen*, *Senator Warren* and *Congressman Ryan*, so that they contained one additional accent.

(1) a. b.



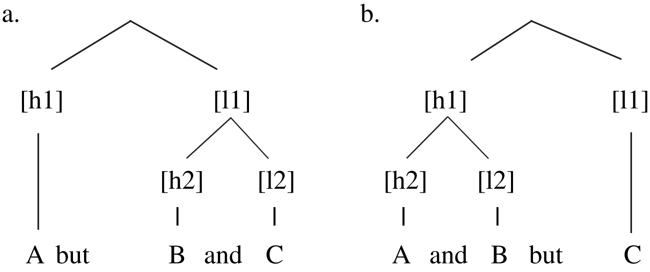
(2) a. Ryan has a lot more money, [but Warren is a stronger campaigner, and Allen has more popular policies]
b. [Allen is a stronger campaigner, and Ryan has more popular policies], but Warren has a lot more money.

Ladd hypothesized that the different syntactic structures induced by the conjunctions *but* and *and* would be reflected in the tonal scaling. Due to the internal complexity of the clauses A, B, and C, such effects would be non-local on the surface.

The hypothesis about structural differences was confirmed. While the high tones in each clause showed clause-internal downstep, the phonetic height of the clauses relative to each other (measured in the relative height of the clause-initial peaks) proved to be sensitive to the hierarchical differences among the two experimental conditions.

Ladd proposed to model his findings by postulating that larger domains such as the constituents A, B, C in (1) enter into downstep interactions relative to each other. The larger domains thus calibrated relative to each other would in turn contain internal downstep. Downstep in this proposal is modeled by abstract register features [h] (high) and [l] (low) assigned to sister nodes, where the left sister is annotated [h] and the right sister is annotated [l]. (3) illustrates for (1) and (2) the relations among the clauses that Ladd proposed.

(3) a. b.



This annotation has phonetic consequences for the scaling of H tones, such that for a sister-pair [h,l], the highest peak in [l] is lowered by one step relative to the highest peak in [h], as expressed by the Relative Height Projection Rule (RHPR) in Ladd (1990:44), reproduced in (4). Ladd modeled this rule on the Relative Prominence Projection Rule of Liberman & Prince (1977), which was aimed at expressing prominence relations among metrical constituents. By the RHPR each pair [h,l] encodes an abstract downstep relation, whereas a pair [l,h] encodes a level relation among the respective highest peaks.

(4) Relative Height Projection Rule (Ladd 1990:44)
In any metrical tree or constituent, the Highest Terminal Element (HTE) of the subconstituent dominated by l is:
a. One register step lower than the HTE of the subconstituent dominated by h when the l subconstituent is on the right.
b. At the same register as the HTE of the subconstituent dominated by h when the l subconstituent is on the left.

According to (4), B in (3a) is lowered relative to A (on account of [h1,l1]), and C is lowered relative to B (on account of [h2,l2]). On the other hand, the representation in (3b) predicts that B is lowered relative to A (on account of [h2,l2]) and that C is also lowered relative to A (on account of [h1,l1]). However, no lowering between B and C is predicted in (3b). This model captures a number of observations concerning the height of the clause-initial peaks in the clauses A, B, and C in Ladd's data, which confirm the Relative Height Projection Rule. In the condition shown in (3a), the first high tone in B is less high than the one in A, and the first high tone in C is less high than the one in B. In the case corresponding to (3b), B starts less high than A, and C also starts less high than A, though importantly, it seems that C is not systematically lowered relative to B. Thus, for some speakers, C is higher than B, for others it is lowered relative to B but less so than in (3a). This account also captures a difference in the comparison of the two clauses (where the first clauses A start at approximately the same phonetic height in the two conditions). The initial peak of C in (3a) (subject to two steps of abstract lowering) is lower than the initial peak of C in (3b) (subject to only one step of abstract lowering).

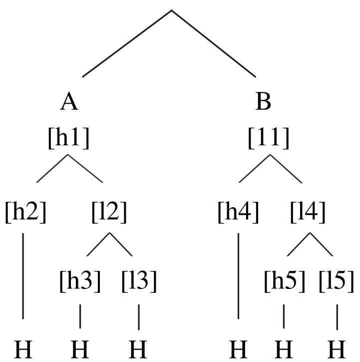
Ladd's data show an additional distinction between the two conditions that is not captured by the Relative Height Projection Rule: Clause B in (3a) begins higher than clause B in (3b). This is not predicted by the hierarchical register representation, in which the initial peak in B should, in both cases, be lowered by one step relative to the initial peak in A. Ladd (1988) subsumes this additional effect under the formulation that in the comparison across the two conditions, a peak is relatively higher after *but* (i.e. after a stronger boundary) than after *and* (i.e. after a weaker boundary), as formulated in (5). This holds of the comparison across the two conditions initially in B, and likewise in the comparison initially in C described above.

(5) Clause-initial accent peaks are higher following a stronger boundary. (Ladd 1988:541)

The effects thus observed are non-local, since the initial peaks of the three clauses are separated by the additional non-initial peaks of the three clauses. Each clause has three peaks in Ladd's first experiment, and four peaks in his second experiment.

The model proposed by Ladd allows the embedding of downstep in smaller domains within downstep in higher domains ('wheels within wheels', as van den Berg et al. 1992 call it). The register representation is applied not only to the two levels of structure displayed in (3), but also to the deeper-embedded clause-internal downstep in these structures, as displayed in (6). Here downstep at the higher level is shown only for a single sister relation between A and B, which stand for clauses. Some amount of downstep internal to A and internal to B is represented in addition.

(6)



This downstep within downstep offers a plausible view of partial reset, a return to a higher register as compared to the immediately preceding high tone. A partial reset is present in position [h4], initial in B in (6). Due to [h4,l4], this is the highest peak within B. ([h5] is lowered relative to [h4] due to [h4,l4], and [l5] is further lowered relative to [h5]). The height of [h4] is a reset because it does not continue the downstep within A that is modeled by [h2,l2] and [h3,l3]. However, it does not return to the initial height found in position [h2]. Rather, [h4] is lowered relative to that initial peak by the downstep relation among A and B, modeled by [h1,l1]. Thus, the tree naturally predicts a partial reset initially in a new larger constituent. It does so by embedding smaller downstep ([h,l]-pairs 2, 3, 4, and 5) within downstep among higher domains ([h1,l1]).

Ladd's experiment can be understood as a way of testing this particular conception of partial reset by nesting partial reset within partial reset. Since the predictions of that model were confirmed in Ladd's results with English speakers (as they were in our experiment with German speakers, see below) the conception of partial reset is empirically supported.

These models may be contrasted with a different way of looking at partial reset: Declination in the sense of Pierrehumbert (1980), rather than downstep, could be hypothesized to account for the lowering among clause-initial peaks. Such declination is constant across the utterance and by hypothesis independent of the structure within the utterance. The prediction of this alternative is therefore that lowering among clause-initial peaks should be identical among the two experimental conditions depicted in (3).

As pointed out by Ladd, Pierrehumbert's view on tonal scaling disallows non-local relationships between tonal domains. Ladd's data can be seen to argue against such an alternative conception of tonal scaling.

2.2. *Phrasal reference line: the experiments of van den Berg et al. (1992)*

The model of embedded downstep has been experimentally strengthened and further developed by van den Berg, Gussenhoven & Rietveld (1992). An illustration of partial reset from this study is reproduced in Figure 1.

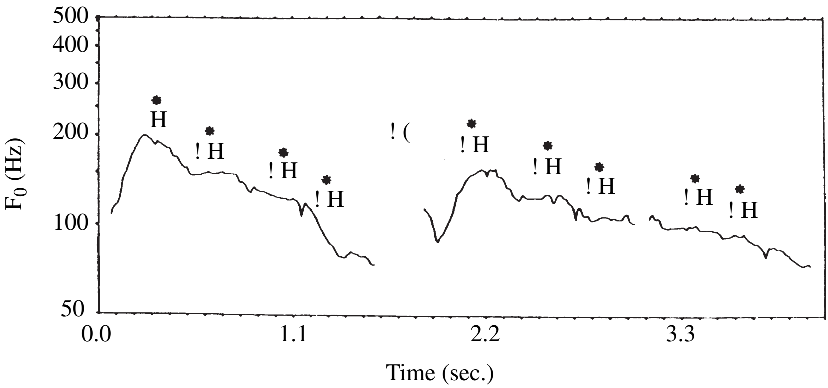
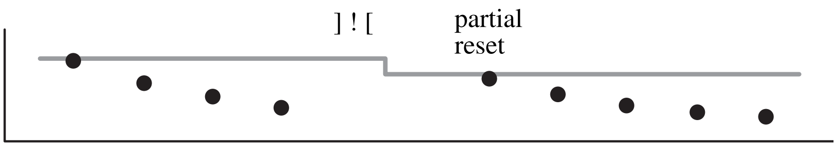


Figure 1. Partial reset in the Dutch utterance (*Merel*, *Nora*, *Leo*, *Remy*), *en* (*Nelie*, *Mary*, *Leendert*, *Mona en Lorna*). From van den Berg, Gussenhoven & Rietveld (1992:334)

Two important findings of this study with Dutch speakers are (a) that reset is a register shift downward relative to a preceding phrase rather than a register shift upward relative to a preceding accent,[1] thus that the nature of the raising process initially in a later domain of a complex utterance is indeed independent of the height of the end of the preceding clause, and (b) that downstep inside a phrase is often larger than downstep across phrases. In other words, downstep among neighboring accents is more dramatic than lowering among reset domains (the latter measured in the height differences among the domain-initial peaks).

Van den Berg et al. (1992) introduce the phonetic abstraction of the phrasal reference line, a register line of constant height during a reset domain (such as a matrix clause), running at the height of the domain-initial peak. Ladd's suggestion of downstep among larger domains is then implemented by the lowering of the phrasal reference line between reset domains, as illustrated in (7).

(7) Partial reset, modeled with a phrasal reference-line (from Truckenbrodt 2002, after van den Berg et al. 1992)



The idea of wheels within wheels is preserved here: smaller downstep (among accents) is embedded in larger downstep (modeled by the phrasal reference line). As in Ladd's account, partial reset results where the effects of smaller downstep are undone, and where the effect of downstep among larger domains can be seen.

2.3. *German clause-final upstep*

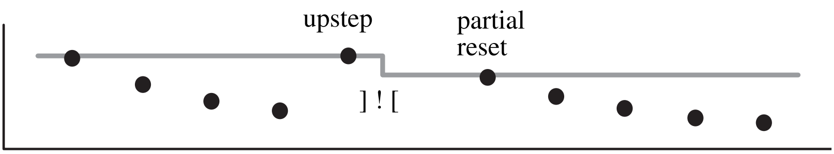
German shows an upstep phenomenon first described by Féry (1993) in connection with focus. In a sequence of two H*(+L) pitch accents, the second is downstepped relative to the first where no narrow focus is involved. However, if the second accent carries narrow focus, it is not downstepped and seems to be scaled to a height comparable to that of the first accent. Upstep is thus understood as a raising of the fundamental frequency back to a preceding high level, following downstep. In English, related observations have earlier been reported in Beckman & Pierrehumbert (1986) in connection with focus, and by Ladd (1983) in what may be independent of narrow focus. The same observation, also in

[1] In analyses using a phonetic register, H tones are scaled to the top of a register interval and L tones to the bottom of the register interval. Changes of tonal height such as downstep are then analyzed as a shift in the height of this register interval – lowering and narrowing, in the case of downstep.

connection with narrow focus, was made for the scaling of successive L* + H accents in Bern Swiss German by Fitzpatrick-Cole (1999). Later Truckenbrodt (2002) found a clause-final upstep independent of narrow focus, and he studied its properties in some detail. This upstep, which seems to be absent in otherwise comparable English data, will be seen in the following to interact with the predictions made by the models of tonal scaling presented in this paper.

In Truckenbrodt's study, speakers from the Southern German-speaking area, including Austria, were recorded. They employed what is analyzed as an L* + H pitch accent on all accents. In prenuclear position, these show downstep from one accent peak to the next. At the end of non-final clauses, some speakers show the phenomenon of upstep on the (L* + H) nuclear pitch accent. Truckenbrodt's analysis and the approximate height-relations of the peaks involved are illustrated in (8). This upstep is arguably independent in its height from preceding downstep, and arguably targets a height comparable to that of the clause-initial peak. It is followed by partial reset initially in the following clause, which appears to be one step of downstep lowered relative to the upstepped peak.

(8) Upstep and partial reset (from Truckenbrodt 2002)



Truckenbrodt (2002) argues that upstep targets the phrasal reference-line of van den Berg et al. (1992) at the end of the first large domain. Upstep thus provides independent evidence for the implementation of embedded domains associated with phrasal reference-lines.

3. Results of Truckenbrodt and Féry (2004)

In an experiment conducted using speakers of Northern German, we wanted to find out whether we also could find the effects of syntactic structure on tonal scaling in stimuli similar to Ladd's sentences. The effects of both downstep and clause-final upstep were investigated. We analyzed data from 5 speakers who read a series of sentences, 16 sentences with the structure A while [B and C], and 16 sentences with the structure [A and B] while C.[2]

[2] The experiment also included 18 longer sentences consisting of three clauses without internal structuring, just separated with a comma and the conjunction *und* 'and' (see Truckenbrodt & Féry 2004 for details). These sentences are marginal for the concern of the present paper.

An example stimulus of each experimental condition is shown in (9) and (10). Expected accent locations are underlined. We call the two experimental conditions AX condition and XC condition. In both AX and XC, the embedded complex clause is abbreviated as X.

(9) AX condition: A while [B and C]$_X$

Warum meint Anna, dass Handwerker teurere Autos haben als Musiker?
Weil der Maler einen Jaguar hat, während die Sängerin einen Lada besitzt, und der Geiger einen Wartburg fährt.
'Why does Anna think that craftsmen have more expensive cars than musicians?
Because the painter has a Jaguar, while the singer possesses a Lada, and the violinist has a Wartburg.'

(10) XC condition: [A and B]$_X$ while C

Warum meint Anna, dass Musiker nicht so teure Autos haben wie Handwerker?
Weil die Sängerin einen Lada besitzt, und der Geiger einen Wartburg fährt, während der Maler einen Jaguar hat.
'Why does Anna think that musicians have less expensive cars than craftsmen?
Because the singer possesses a Lada, and the violinist drives a Wartburg, while the painter has a Jaguar.'

The tonal contours of these sentences that were most frequently found are illustrated in (11). For the purposes of this paper, we analyze each clause as a prosodic intermediate phrase ip (Beckman & Pierrehumbert 1986), and the combination of two clauses to a higher constituent as an intonation phrase IP. The non-final ips consist of rising tones L* + H and a boundary tone, transcribed as H- here. The final clause has one rising pitch accent and one that is sometimes falling, sometimes a single tone, and ends low.

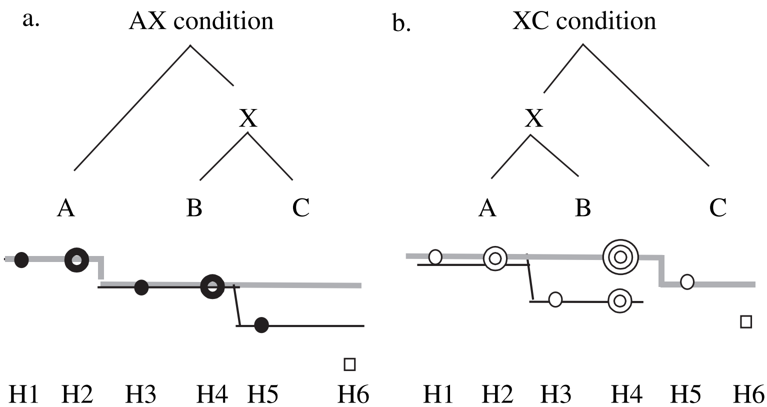
(11) Warum meint Anna, dass Sportler nicht so teure Autos haben wie Handwerker?

L* + H1 L* + H2 H-
[Weil der Ringer einen Lada besitzt]$_{ip}$
L* + H3 L* + H4 H-
[während der Maler einen Jaguar fährt]$_{ip}$ } IP
L* + H5 (H6 +)L* L-L%
[und der Weber einen Daimler hat]$_{ip}$

An alternative representation would choose the prosodic level IP for the prosodic constituents around the individual clauses, rather than ip. Retaining the higher IP constituent shown in (11), this would then amount to recursive IPs as suggested by Ladd (1986). We believe that there is no evidence in our material for choosing between ip and IP (and thus also no real evidence for an ip in German). We choose these labels for concreteness here.

(12) shows a schematic phonetic analysis of the two structures, using the phonetic reference-lines of van den Berg et al. (1992). Importantly, the additional internal constituents $[BC]_X$ in (a) and $[AB]_X$ in (b) are also represented by phrasal reference lines (the thick grey reference line in (12)). This is crucial in (b), where C is downstepped relative to the reference line of the complex constituent $[AB]_X$. Expectations about accent peak height and upstep are shown by the dots and circles. For the initial peaks in A, B and C (H1, H3, H5, the dots), these reflect what was discussed in connection with Ladd's experiment and proposal above. In addition, upstep would be expected in clause-final position of A and B (H2, H4, the circles). Recall that upstep of this kind does not normally occur in utterance-final position. It is thus not expected at the end of C (H6, the squares). The height of the upstepped peaks H2 and H4 is expected to be comparable to the initial peak height in the same clause in A and B in (a) as well as in clause A in (b). Of particular interest, however, is the scaling of the upstepped peak H4 in clause B in (b). One possibility is that this upstepped peak is scaled to the phrasal reference-line of B, downstepped relative to A, at the same height as the initial peak in B (the circle). A second possibility is that the scaling of upstep at the end of B targets the combined phrasal reference-line of [AB], at the height of the initial peak in A (the triple circle). This would be detectable in the comparison of the upstepped peak at the end of B with the clause-initial peaks of A and of B. It would further be expected to contrast with (a), where the analysis does not provide such an option of higher upstep at the end of B.

(12)



Our empirical results on clause-initial peaks replicated the results of Ladd (1988) reviewed above and thus confirmed the predictions made by the model. First, as shown in Fig. 2, in the AX condition, there is continuous downstep across the three clauses in the clause-initial peaks, H3 being lowered relative to H1 and H5 relative to H3. By contrast, in the XC condition, steep downstep is observed between H1 and H3, but none between H3 and H5. This is in line with Ladd's results, and is consistent with the predictions of his model.

Second, the effect of (5) above: The third high tone (H3) is higher in the AX condition than in the XC condition. This can also be seen in Fig. 2 which compares the average height of the clause-initial high tones in our two conditions. We were thus able to replicate Ladd's findings about the influence of hierarchical structure on the tonal phrasing at the level of the clauses.

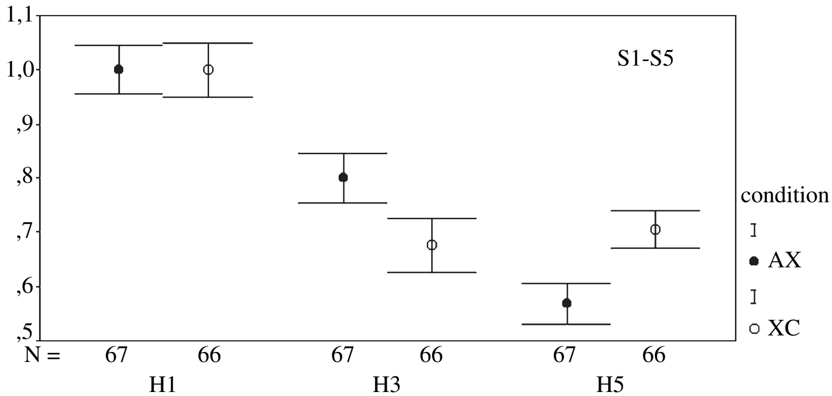


Figure 2. 95% confidence intervals of the clause-initial peaks in the pooled normalized values of the two experimental conditions. From Truckenbrodt & Féry (2004)

In the non-final clauses, the expected upstep was observed (see Truckenbrodt & Féry 2004 for more details). In the AX condition (compare (12a)), H2 was of comparable height to H1, and H4 was of comparable height to H3, with both H3 and H4 lowered relative to H1 and H2. In the XC-condition, (compare (12b)), two groups of speakers are distinguished. The group of interest here consists of the three speakers S3–S5, whose normalized values are displayed in Figure 3, preserving plotting style relative to (12b). It can be seen that H2 is of comparable height to H1, due to upstep, as expected. Interestingly, for these three speakers, H4 is not of comparable height to H3. Instead, H4 is significantly higher than H3, approximating H1 in height. We interpret this as evidence that these three speakers target the thick grey reference-line of (12b) in H4, the reference-line of the crucial constituent X (comprising A and B).

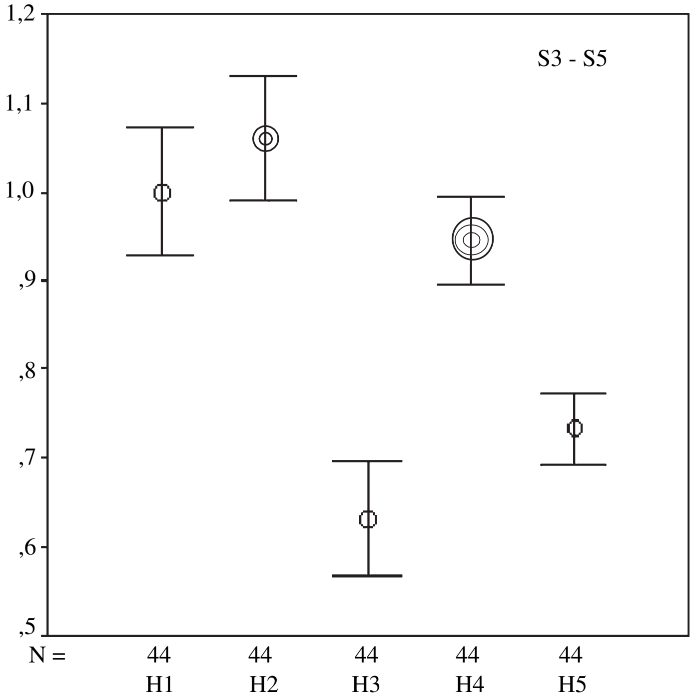


Figure 3. 95% confidence intervals for upstep in clause B of the XC condition (H4) in the normalized values of three speakers. From Truckenbrodt & Féry (2004)

As pointed out in Truckenbrodt & Féry (2004), the realization of this pattern of scaling by some speakers is interesting evidence, not only for the phrasal reference-lines of van den Berg, Gusssenhoven & Rietveld (1992) and the scaling of upstep according to Truckenbrodt (2002), but also for Ladd's account of the distinction between the two experimental conditions in terms of hierarchically conditioned downstep relations. Thus, in Ladd's account, the important distinction between the two experimental conditions is observed in the relative height initially in C. While C is downstepped relative to B in the AX condition, C is crucially downstepped relative to $[A \& B]_X$ in the XC condition (see (12b)). In our data, speakers S3–S5 seem to upstep H4 to precisely the reference-line of this crucial constituent $[A \& B]_X$ in the XC condition. The upstep of H4 for these speakers thus provides independent evidence of the presence of the reference-line of the complex constituent $[A\ \&\ B]_X$, relative to which constituent C is crucially downstepped in Ladd's account.[3]

4. Sisterhood in layered structure and tonal scaling

In this section we describe our theoretical model. In the model, the constituents to which tonal scaling refer are prosodic domains rather than morpho-syntactic constituents. Sisterhood among adjacent constituents is interpreted as register lowering. The new idea here is the suggestion that the more deeply embedded the constituents are, the steeper the downstep.

4.1. *N-ary prosodic constituency*

We begin this section by backtracking a bit in the history of the development of ideas. As mentioned in section 2.1, Ladd explicitly suggested a hierarchical representation of register in parallel to the hierarchical representation of prominence developed by Liberman & Prince (1977). In the latter proposal, the hierarchy was given by the morpho-syntactic structure of the elements involved (as in Cooper & Sorensen 1981). However, subsequent work has argued that hierarchical assignment of prominence is relative to prosodic categories with partial similarities to morpho-syntactic categories, rather than to morpho-syntactic structure directly (see Selkirk 1980 for the foot and the prosodic word, and Nespor & Vogel 1986 for higher prosodic constituents).[4] Here we offer a model in which the hierarchical structure relevant for tonal register is likewise prosodic in nature.

An important distinction between syntactic and prosodic structure is that prosodic structure is by default non-recursive (no foot inside another foot, for example), while no ban seems to exist on recursion in syntactic structure (verb phrases inside of verb phrases, for example). This distinction lead to some shifts in the analysis of prominence, when the hierarchical syntactic theory of prominence was changed to a

[3] One might wonder whether the differences in height of the clause-initial peaks across the two experimental conditions would also be compatible with an interpretation in which the juxtaposing or contrasting meaning of *während* 'while' triggers a following boosting effect. Such a boosting effect, rather than boundary strength difference, could be held responsible for the difference between experimental conditions. See also Ladd (1988:540) on this issue in his materials. We believe that the finding in connection with upstep is interesting evidence against this interpretation: H4 is comparable in height to H3 in the AX condition, but comparable in height to H1 (for S3–S5) in the XC condition. This difference follows from the interpretation in the text, but not from the analysis in terms of boosting after *während.*

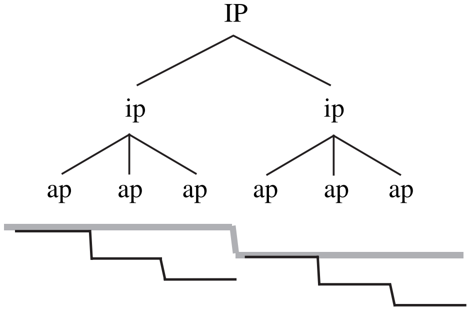
[4] Thorsen's (1980) results for Danish likewise show that prosodic phrases are not isomorphic with syntactic constituents. Thorsen conducted experiments to investigate the effect of length of utterances on the scaling of initial accents, and arrived at the conclusion that there is an effect of the utterance's length on the scaling of initial accents, though she could not offer a generalization for her results due to lack of consistent effects.

hierarchical prosodic theory of prominence. We suggest a similar shift in the hierarchical representation of register. We use the hierarchical structure proposed by Beckman & Pierrehumbert (1986) and Pierrehumbert & Beckman (1988) who use intonation phrases (IP), intermediate phrases (ip) and accentual phrases (ap).[5] We assume a provision by which each syntactic clause forms either an intermediate phrase or an intonation phrase. This is compatible with the non-recursive phrasings $[ip]_{IP}$ $[ip\ ip]_{IP}$ as well as $[ip\ ip]_{IP}$ $[ip]_{IP}$ in the two experimental conditions, with prosodic structure mirroring syntactic structure. We further adopt the notion of van den Berg et al. (1992) that certain higher prosodic constituents correspond to a particular register level and generalize this to prosodic categories more generally (see also Pierrehumbert & Beckman 1988). With this, we employ the two principles in (13) to relate hierarchical prosodic structure to register scaling, following up on the suggestions in Truckenbrodt & Féry (2004).

(13) a. First sister
The leftmost daughter of a node shares the register properties of its mother.
b. Little sister
Prosodic sisterhood among adjacent constituents is, by default, phonetically interpreted as register lowering.

An important difference between morpho-syntactic and prosodic domains concerns the n-arity of the sisterhood. Ladd's representations derive a step of lowering among sister nodes through the assignment of [h,l] to binary structures. As far as binary patterns are concerned, our modification is equivalent to his account, deriving the lowering relation from the configuration of sisterhood by (13b). However, (13b) extends this lowering relation to multiple sisters. The extension is exemplified in (14), where each prosodic sister relation is now interpreted as downstep.

(14)



[5] Whether or not this structure may carry over to English depends on the assignment of edge tones in sentences of the kind used by Ladd.

The prosodic foundation of register relations allows us to overcome a criticism of Ladd's theory by van den Berg et al. (1992). Ladd's theory leads to the expectation that clause-internal downstep also depends on the hierarchical (syntactic) organization of the constituents. This aspect of Ladd's theory is criticized by van den Berg et al. (1992), who maintain that clause-internal downstep is independent of the syntactic configuration in which the constituents occur. The issue also arises with respect to the material in Truckenbrodt (2004), where prosodic sisterhood rather than syntactic sisterhood can be argued to be the downstep triggering configuration. Thus, syntactic hierarchical relations *among clauses* may lead, as in Ladd's experiment, to hierarchical register reflexes. *Clause internally*, however, hierarchical relations do not seem to affect the pattern of downstep. Rather, where downstep applies, each accent seems to be downstepped relative to the preceding one, regardless of their structure, as maintained by van den Berg et al. (1992). This clause-internal downstepping behavior can be derived from (11) in a prosodic representation in which a series of accentual phrases are prosodic sisters to each other.

We believe that it is reasonable to extend the conclusion that prosody, rather than syntax, triggers downstep among accentual phrases to the higher levels of ip and IP. We suggest that the two experimental conditions are not only syntactically distinguished, but prosodically as well, with prosody mirroring syntax.

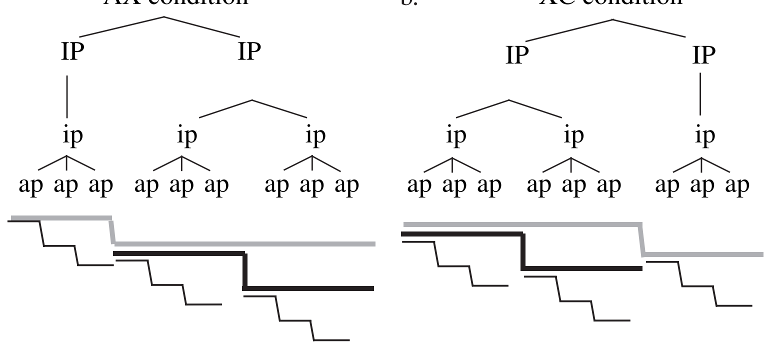
4.2. *Hierarchical structure and reference lines*

As shown in (14), 'First sister' and 'Little sister' in (13) not only translate prosodic sisterhood into downstep, but assign the phonetic reference lines to prosodic levels and define the course of multiple simultaneous reference lines. If each reference line is associated with a particular prosodic level, then (13) correctly separates the different prosodic levels, and correctly assigns their relative phonetic height. The phrasal reference line (the grey line in (14)) is seen as a property of the larger domains, here of the ips in (14). This reference-line is lowered between the two sister ips on account of (13b). At the same time, the black reference lines in (14) stem from the structure of the accentual phrases. Here (13a) identifies the height of the first ap with that of the ip above it. (13b) then lowers the reference-level for each further ap in the same ip relative to its preceding ap sister.

Notice that our minimal modifications allow us to retain the important insight that downstep may be embedded inside of downstepped constituents. In the simplest case, this amounts to the analysis of the partial reset as a return to an independent register (thus it is a reset) which may be downstepped at a higher level (it is thus partial). In (14), such a partial reset is seen in the fourth accentual phrase (ap). It shows the effect of lowering among the two ip sisters. However, lacking preceding ap sisters, this ap does not show any further register subordination.

The multiply embedded cases, schematically shown in (15), involve one more level of complexity. In Ladd's English data, these feature two partial resets, initially in the second and third clause. As was seen both in Ladd's findings and in ours, the clause-initial peaks depend for their height on the syntactic structure of the three clauses. In the slightly revised analysis suggested here, these hierarchical syntactic differences translate into isomorphic prosodic differences as shown, from which the register relations follow by (13). In (15a), the second IP corresponds to the complex constituent X. The reference-level is lowered between the two IP sisters (thick grey line), which plays out initially in the second IP, where clause B is lowered relative to clause A. Clause C is then further lowered relative to B due to the sister-relation among the two ips (thick black line). In (15b), the complex constituent X corresponds to the first IP. This crucially plays out in the scaling: Lowering among the two IPs amounts to lowering of C relative to [A & B]$_X$ (thick grey line). In addition, B is lowered relative to A due to the sister-relation among the two ips (thick black line). Thus, our modification can represent the difference between the two experimental conditions, while still allowing embedded clause-internal downstep, in a way that mirrors Ladd's original suggestions.

(15) a. AX condition b. XC condition

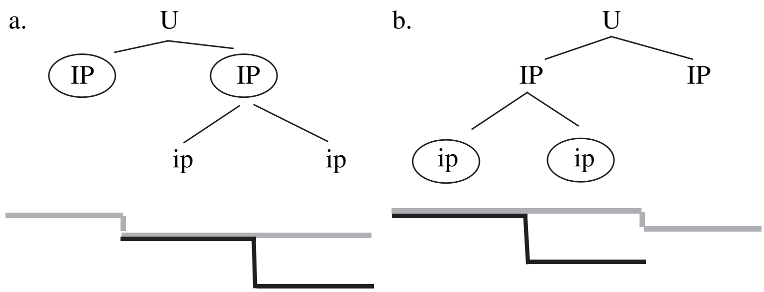


4.3. *The consequences of embedding on tonal scaling*

In Ladd's data, as well as in our reproduction of them, the initial peak in the second clause is downstepped more in (15b) than in (15a). This difference is not accounted for in either Ladd's hierarchical representation or our modification of it. Recall that Ladd accounts for this difference by referring to the strength of syntactic boundaries, as formulated in (5). In (16), we here propose a revision of this hypothesis that incorporates (5) and finds independent support elsewhere. (16) is compatible both with Ladd's proposal in terms of register features and with our modification of it. Instead of boundary strength, it makes reference to the related notion of levels of embedding in the prosodic representation.

(16) The deeper the steeper (TDTS)
Downstep among sister-nodes is relatively stronger for constituents relatively lower in the hierarchical representation.

The application of TDTS that is shared with (5) is shown in (17).

(17)


In (17a), the downstep relation that affects the second clause relative to the first reflects the sisterhood relation of the two IPs, immediately below the root node U. By contrast, in (17b), the downstep relation between the first and the second clause reflects the sisterhood relation between the two ips; these are joined into an IP that joins with another IP to form the utterance U. The ip sisters in (17b) are thus one level lower in the hierarchical representation than the IP sisters in (17a). (16) captures the fact that the lesser embedding in (17a) leads to quantitatively more shallow downstep between the first and second clause, while the deeper embedding in (17b) leads to more dramatic downstep between the first and second clause. In this regard, TDTS mirrors Ladd's formulation in (5): Less embedding of the circled constituents in (17a) than in (17b) is comparable to Ladd's stronger boundary between the circled constituents in (17a) than in (17b).

Independent support for the connection postulated in (16) comes from the result of van den Berg et al. (1992) mentioned above. They find that the ratio of downstep is more dramatic among adjacent accents than among larger domains (such as from the initial peak to the first partial reset). This observation can be subsumed under (16):

the accentual phrases that are prosodic sisters and would define downstep among adjacent accents are structurally more deeply embedded than the adjacent larger domains that contain them, and would correspondingly show stronger downstep than the larger domains. Notice that this case cannot be subsumed under (5), which allows comparison only of different resets, and not of accentual downstep with phrasal downstep.

Figure 4 compares different downstep ratios from our data, calculated against the height of the utterance-final L% tone value of a speaker. The values plotted are ratios of downstepped H tones (above L%) to earlier H tones (above L%). By and large the values corroborate the prediction about the relationship between embedding and downstep.

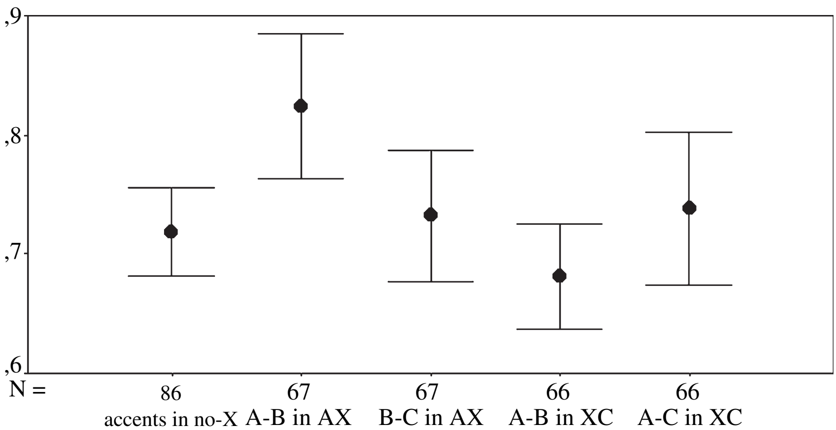


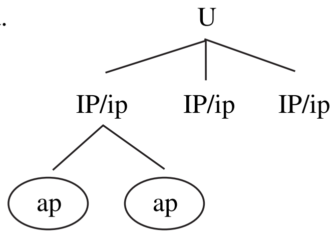
Figure 4. 95% confidence intervals for different downstep ratios on the basis of the pooled normalized data of the five speakers of Truckenbrodt & Féry (2004)

The first value on the left is from a part of the experiment that we call "no-X"-condition (see footnote 1). It reflects clause-internal downstep. This value is a bit higher than 0.7, meaning that the height of the second accent (above the value of L%) is on average about 0.7 of the height of the first accent (above the value of L%). This is here compared with the downstep ratios among the clauses in the AX and XC conditions, assessed in the clause-initial peaks. From left to right, downstep among A and B in the AX condition, between B and C in the AX condition, between A and B in the XC condition, and between A (not B!) and C in the XC condition.

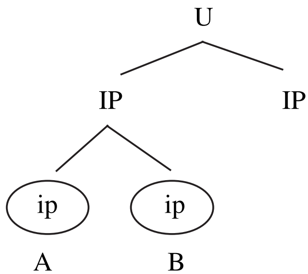
Downstep among A and B in the AX condition (around 0.83) is less strong than downstep among the first two accents in the no-X condition. This corresponds to the observation that clause-internal downstep, as in the no-X condition, is stronger than downstep across the clauses A and B in the AX condition, which takes place between two IPs. This reproduces the observation of van den Berg et al. (1992) that downstep among higher domains is more shallow than downstep among accents, and is captured by TDTS, as was seen above.

Downstep among A and B in the XC condition (around 0.68) is much stronger than downstep between A and B in the AX condition. This is the distinction across the two conditions initially in clause B that was illustrated in (17). It is derived by Ladd's formulation in (5), as well as by TDTS, as was seen. What is also interesting here is that downstep among A and B in the XC condition is comparable to clause-internal downstep among the first two accents in the no-X condition. This is compatible with TDTS, as shown in (18ab).

(18) a.



b.



This case suggests two things. For one thing, that 'only child embedding' is ignored by TDTS. By 'only child embedding', we mean the embedding of a constituent like ip in a higher constituent like IP, though without any sister ip nodes inside of the higher IP. Otherwise, the two layers IP and ip would add to the level of embedding in (18a) and predict steeper downstep there than in (18b). However, Figure 4 shows that the degrees of lowering are comparable (or, if anything, the inverse of this expectation). The other conclusion suggested by this case is that the label of the prosodic level does not matter for TDTS. If it did, the lower level of aps in (18a) should show more downstep than the higher level of ip in (18b), again contrary to the results.

The remaining values, involving clause C, are broadly compatible with this picture, though less clearly set off in the expected directions. Downstep between B and C in the AX condition is embedded in [B & C]$_X$, and thus has two levels of embedding, like the cases in (18). Figure 4 suggests that the amount of downstep here is indeed comparable to these two other cases. Finally, downstep between A and C in the XC condition is expected to be less steep than found in Figure 4, with only one level of embedding. A possible explanation for the discrepancy between actual and expected values here is that declination affects this last value more than the others. Downstep between A and C may show overlaid effects of declination more than the other cases, since it is assessed across an intervening clause, while the remaining values assess downstep among adjacent clauses.

It is also tempting to apply the principle TDTS to the Japanese observation of boosting by Kubozono (1989). In Kubozono's core case, he compares the structures [[XY]Z] and [X[YZ]], where X, Y, and Z are all accented elements in the same major phrase. In both cases, there is downstep from each accent to the next. However, the downstep between X and Y is stronger in [[XY]Z] than it is in [X[YZ]]. This latter observation can be made to follow from TDTS if it is assumed that the syntactic structure is here mirrored by recursive minor phrase structure: The sister-pair X and Y is one branching node removed from the root node in [[XY]Z], while X and [YZ] are immediate daughters of the root node in [X[YZ]]. With deeper embedding in the former case, steeper downstep is correctly predicted. However, if the intermediate constituent [XY] is prosodically represented in [[XY]Z], we might expect that Z is not downstepped relative to Y, much as in the XC condition of Ladd's experiment and ours (cf. (17b)). This is contrary to Kubozono's findings, where Z is downstepped relative to Y in both conditions. In other words, the observation can be derived if the intermediate constituent [XY] is present for the purpose of TDTS, but nevertheless does not introduce a reference-line relative to which the following constituent Z is scaled by lowering. There are a variety of hypotheses that could achieve this. The one that we tentatively offer here is the following: Assume that only certain prosodic constituents introduce a reference-line relative to which later constituents can be scaled by lowering (i.e. that there are only certain prosodic constituents relevant to (13)): those that have a prosodic head that marks their strongest element. In Japanese, these may be accented minor phrases (with the accent marking the head of prominence) as well as major phrases (with the initial minor phrase being the head). This would correctly exempt the additional higher minor phrase [XY] which represents a grouping but may be said to have no separate expression of prominence. It would also correctly exempt Japanese unaccented minor phrases from participating in the downstepping pattern defined by (13).

Before concluding, we point out that there are cases of conditions on downstep (other than the difference between unaccented and accented minor phrases in Japanese) that ultimately need to be integrated into a comprehensive understanding of downstep. In some languages, certain tonal configurations are required for downstep, such as the presence of L tone for downstep to take effect in Yoruba (Connell & Ladd 1990, Laniran 1992) or the possible absence of downstep in sequences of H* in English, in contrast to the downstepping binary pitch accents in the analysis of Beckman & Pierrehumbert (1986). Also in German, downstep is absent from certain hat patterns related to information structural contents, as shown in Féry (1993) and Féry & Hartmann (2004). It appears that downstep can furthermore be blocked in certain semantic contexts. Examples of this can be found in the observations of Ladd and van den Berg et al. (1992). It has been suggested by Ladd (1983) and van den Berg et al. (1992) that downstep is represented by a phonological annotation or feature. That approach will allow for different representations in the downstepping and non-downstepping cases. Our suggestion about the correlation between hierarchical structure and downstep above does not take these additional observations into account. Rather, it seeks to isolate the contribution of hierarchical structure to downstep. Clearly, this can only be one element in a more complex, integrated theory of downstep, in which these other observations must likewise find their place.

5. Conclusion

We have reviewed results and suggestions by Ladd (1988), van den Berg et al. (1992), Truckenbrodt (2002), and Truckenbrodt & Féry (2004) that relate hierarchical structure and tonal scaling. We have emphasized the relevance of prosodic constituents. Following van den Berg et al. (1992), these correspond to register levels on which clause-initial (and, according to Truckenbrodt 2002, upstepped) tones are scaled. In these terms, an analysis is possible that preserves what we view as the important insights of Ladd's (1988) register-feature representations. We have also argued for an addition to the theory of down step, the principle 'The Deeper, the Steeper'. By this principle, the level of embedding, counting only branching constituents, determines the strength of lowering between two sister-nodes.

References

BECKMAN, M. E. & PIERREHUMBERT, J. B. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3, 255–309.
CONNELL, B. & LADD, D. R. 1990. Aspects of pitch realization in Yoruba. *Phonology* 7, 1–29.
COOPER, W. & SORENSEN, J. 1981. *Fundamental frequency in sentence production*. Heidelberg. New York: Springer.
FÉRY, C. 1992. Focus, topic and intonation in German. *Arbeitspapiere des Sonderforschungsbereichs 340*, 20. University of Tübingen.
FÉRY, C. 1993. *German intonational patterns*. Tübingen: Niemeyer.
FÉRY, C. & HARTMANN, K. 2005. Focus and prosodic structure of German gapping and right node raising. *The Linguistic Review* 22, 67–114.
FITZPATRICK-COLE, J. 1999. The alpine intonation of Bern Swiss German. *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, 941–944.
KUBOZONO, H. 1989. Syntactic and rhythmic effects of downstep in Japanese. *Phonology* 6, 39–67.
LADD, D. R. 1983. Phonological features of intonational peaks. *Language* 59, 721–759.

Ladd, D. R. 1986. Intonational phrasing: The case for recursive prosodic structure. *phonology yearbook* 3, 311–340.

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

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

Laniran, Y. 1992. *Intonation in tone languages: The phonetic implementation of tones in Yorùbá*. Cornell University. PhD diss.

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

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

Pierrehumbert, J. B. 1980. *The phonology and phonetics of English intonation*. MIT. PhD diss.

Pierrehumbert, J. B. & Beckman, M. E. 1988. *Japanese tone structure*. Cambridge, MA: MIT Press.

Selkirk, E. O. 1980. The role of prosodic categories in English word stress. *Linguistic Inquiry* 11, 563–605.

Thorsen, N. 1980. Intonation patterns and stress group patterns in declarative sentences of varying length in ASC Danish. *Annual Report of the Institute of Phonetics*, University of Copenhagen, vol. 14, 1–29.

Truckenbrodt, H. 2002. Upstep and embedded register levels. *Phonology* 19, 77–120.

Truckenbrodt, H. 2004. Final lowering in non-final position. *Journal of Phonetics* 32, 313–348.

Truckenbrodt, H. To appear. Upstep on edge tones and on nuclear accents. In *Tone and intonation* (working title). eds. C. Gussenhoven & T. Riad. Berlin: Mouton de Gruyter.

Truckenbrodt, H. & Féry, C. 2004. *More on hierarchical organization and tonal scaling*. Ms. University of Potsdam and University of Tübingen. [Intended for publication in Phonology].

van den Berg, R., Gussenhoven, C. & Rietveld, T. 1992. Downstep in Dutch: Implications for a model. In *Papers in laboratory phonology II: Gesture, segment, prosody*. eds. G. J. Docherty & D. R. Ladd, 335–367. Cambridge: Cambridge University Press.

Caroline Féry
University of Potsdam
Institute of Linguistics
PO 601553
DE-14415 Potsdam
Germany
fery@rz.uni-potsdam.de

Hubert Truckenbrodt
Seminar für Sprachwissenschaft
Universität Tübingen
wilhelmstraße 19
DE-72074 Tübingen
Germany
hubertt@uni-tuebingen.de