

# *Hierarchical organisation and tonal scaling\**

**Hubert Truckenbrodt**

Centre for General Linguistics (ZAS), Berlin and Humboldt University, Berlin

**Caroline Féry**

Goethe University, Frankfurt

---

Ladd (1988) investigates configurations A[BC] *vs.* [AB]C of three English clauses containing clause-internal downstep. A sister-node relation between clauses (but not a sequential relation) leads to downstep among clauses, such that C is systematically lower than B in A[BC] but not in [AB]C. These findings are replicated here with German data. In addition, the German phenomenon of upstep (Truckenbrodt 2007b) arguably targets the phonetic reference line that models lowering among clauses (van den Berg *et al.* 1992). We show that both Ladd's and our results also support Selkirk's (2011) suggestion that root sentences/illo-cutionary clauses can be interpreted as matched to intonational phrases (not just aligned with them, as in Downing 1970). The results also suggest that, in addition to downstep among intonational phrases, phrase-final lowering takes place.

---

## 1 Introduction

### 1.1 Overview

In this paper we present and discuss the results of an experimental study on German intonation. A number of related goals are pursued.

Firstly, we replicate the core results for English of Ladd's (1988) seminal paper. These results rest on a comparison of two different structural configurations among three clauses (domains of partial F<sub>0</sub> resetting), i.e. A, B and C, one assuming the structure [A[BC]], and the other the structure [[AB]C]. The replication strengthens Ladd's core conclusions about the effects of hierarchical structure on tonal scaling.

\* E-mail: [HUBERT.TRUCKENBRODT@GMAIL.COM](mailto:HUBERT.TRUCKENBRODT@GMAIL.COM), [CAROLINE.FERY@GMAIL.COM](mailto:CAROLINE.FERY@GMAIL.COM).

We would like to thank Bob Ladd and an anonymous reviewer for helpful comments, and Kirsten Brock for proofreading and stylistic work on this article. All errors are of course our own. The research reported here was made possible by a guest professorship at the Centre of Cognitive Studies in Potsdam during the winter semester 2002–03 for the first author. This research was supported by the SFB 632 on Information Structure at the University of Potsdam for the second author, and by the German Federal Ministry of Education and Research (Grant 01UG1411).

Another goal relates to a phenomenon of clause-final upstep shown by many German speakers. Truckenbrodt (2002, 2007b) argues that upstep supports a particular implementation of Ladd's ideas suggested by van den Berg *et al.* (1992). We are interested here in studying the interaction of upstep with the clausal configurations [A[BC]] and [[AB]C]. Our observations will allow us to further strengthen Ladd's conclusions and their implementation by van den Berg *et al.* (1992).

The results of the study are relevant to the mapping between syntax and prosody. The syntactic clause structures [A[BC]] and [[AB]C] must be mapped to isomorphic prosodic structures if they are to be able to affect the intonation. Reinforcing Ladd's conclusion about the different intonation of the two structures allows us to argue that the Match theory of Selkirk (2009, 2011) and the recursive model of prosodic structure proposed by Féry (2010, 2011, 2015) fare better empirically than an account in terms of alignment and wrapping (Selkirk 2005, Truckenbrodt 2005).<sup>1</sup>

The remainder of this introduction reviews relevant background. §2 introduces the experimental method employed, and §3 presents the results. In §4 we argue for a prosodic distinction between the two experimental conditions, and discuss consequences for the syntax–phonology mapping. §5 sums up our conclusions. Additional details of our method are given in the Appendix.

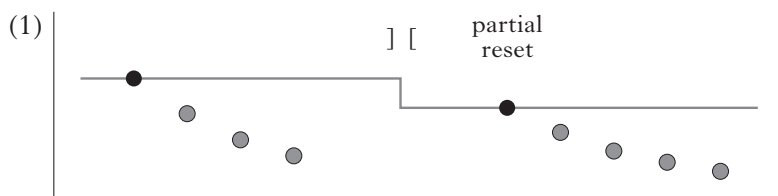
## 1.2 Background on hierarchical organisation and tonal scaling

This paper adopts the autosegmental-metrical analysis of intonation developed in Pierrehumbert (1980), Beckman & Pierrehumbert (1986), Pierrehumbert & Beckman (1988) and later work (see Gussenhoven 2004 and Ladd 2008, as well as papers in Jun 2005, 2014, for recent cross-linguistic overviews and discussion). The F<sub>0</sub> contour of an utterance is analysed in terms of H and L tones. In intonation languages, these belong either to pitch accents (H\*, L\*+H, etc.) or edge tones (here H%, L%). Turning points in the F<sub>0</sub> contour are taken to be evidence for the presence of H or L. Underlying the F<sub>0</sub> contour is a language-specific tonal system, with specific phonological properties and pragmatic meanings (see e.g. Pierrehumbert & Hirschberg 1990, Bartels 1999, Gussenhoven 2000 and Truckenbrodt 2012), as well as phonetic implementation (see e.g. Bruce 1977, Liberman & Pierrehumbert 1984, Pierrehumbert & Beckman 1988 and Ladd & Schepman 2003). The present study is concerned with the phonetic implementation of the tones, in particular the assignment of phonetic values to H tones.

Background for our discussion is the phenomenon of DOWNSTEP of successive H tones, i.e. the successive lowering of the phonetic values of

<sup>1</sup> The material presented here overlaps with Féry & Truckenbrodt (2005), a paper that builds on an earlier version of the current article. The way in which final lowering among intonational phrases strengthens the conclusions about the crucial intermediate constituent (§3.2), and the theoretical conclusions concerning the syntax–phonology mapping (§4), seem to us important enough to publish in a revised version of this earlier manuscript, despite the time that has passed.

the H tones under language-specific phonological conditions.<sup>2</sup> Initially in a new prosodic domain, such downstep can be undone by what is sometimes called a RESET, frequently a PARTIAL RESET (Ladd 1988, van den Berg *et al.* 1992, Truckenbrodt 2002, 2007b, Laniran & Clements 2003), which represents an incomplete return to the utterance-initial F<sub>0</sub> height. We are here concerned with the suggestion of Ladd (1988, 1990) that partial reset involves lowering of one large domain relative to another such larger domain. Implemented in terms of a horizontal reference line that is lowered between the domains, as in van den Berg *et al.* (1992), this takes the form shown in (1). Downstep among accentual H tones, represented by the circles in (1), proceeds away from the phrasal reference line in each domain. In this section the circles representing phonetic H-tone values on the reference line are black and those representing H-tones that are below the reference line as a result of downstep are grey. The reset in the second domain is partial rather than complete, because the phrasal reference line to which the reset returns is itself lowered between the two domains. This lowering of the reference line is also referred to as downstep, though it is not downstep among accents, but rather among the reference lines of two large domains.



The specific combination of these earlier suggestions that we employ is that prosodic sisterhood in the phonology corresponds to a lowering of the reference line in the phonetics. Evidence for the connection of sister nodes to domain-lowering is provided by Ladd (1988). The stimuli of Ladd's experiment can be described as involving the embedding of one partial reset among sister nodes within another partial reset among sister nodes. The sentences in (2) illustrate his two experimental conditions.

(2) a. *but/and* condition: [A *but* [B *and* C]<sub>X</sub>]

Ryan has a lot more money but [Warren is a stronger  
campaigner and Allen has more popular policies]<sub>X</sub>

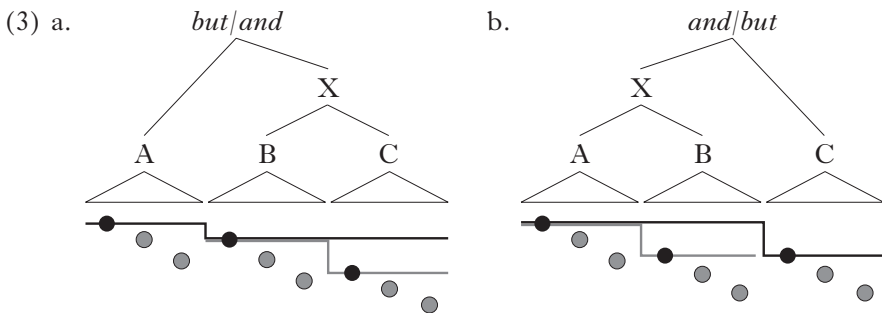
b. *and/but* condition: [[A *and* B]<sub>X</sub> *but* C]

[Allan is a stronger campaigner and Ryan has more popular  
policies]<sub>X</sub> but Warren has a lot more money

<sup>2</sup> Experimental studies of downstep include Liberman & Pierrehumbert (1984) on English, Pierrehumbert & Beckman (1988) on Japanese, Prieto *et al.* (1996) on Mexican Spanish and Laniran & Clements (2003) on Yoruba. A discussion of different downstep triggers in different African languages can be found in Odden (1995).

Both experimental conditions consist of three clauses, A, B and C. In the first condition [B and C]<sub>X</sub> form an embedded constituent, in the second condition [A and B]<sub>X</sub>. The embedded constituent thus formed, here labelled X, is then joined with the remaining clause on the highest level by the conjunction *but*, giving [A but X] in the *but/and* condition, and [X but C] in the *and/but* condition.

Ladd's lowering among sister nodes, modelled in terms of the phonetic reference lines, is illustrated in (3) for these more complex cases. Intuitively, it can be thought of as two applications in each condition of the schema in (1). First, within X, we find lowering of the grey reference line in (3) between the two clauses that are sisters inside X (i.e. between B and C in the *but/and* condition and between A and B in the *and/but* condition). Second, the constituent X is also assigned a reference line of constant height. The sisterhood relation between X and its sister then leads to a second application of the schema in (1), so that there is lowering between X and its sister node (i.e. between A and X in the *but/and* condition and between X and C in the *and/but* condition). This is shown by the black reference line, the grey reference line and the circles, combine as follows: the phonetic height of the leftmost element of one level is defined by the reference line of the next higher level. Thus, in the *but/and* condition in (3a), the height of B, which is initial in the higher X, is defined by the height of X. In the *and/but* condition in (3b), the height of A, which is initial in X, is defined by the height of X. In each clause, the height of the leftmost accent is defined by the height of the reference line of that clause.



The predictions of this model can be assessed with respect to the height of the clause-initial peaks. As shown, in both conditions the first peak in B is predicted to be lowered relative to the first peak in A. This is due to the sister relation [A X] in the *but/and* condition in (3a), and the sister relation [A B] in the *and/but* condition in (3b). Furthermore, in the *but/and* condition, B and C are sisters inside X. Lowering between these sisters predicts that C is further lowered relative to B. The prediction for the *but/and* condition is therefore successive lowering of the clause-initial peaks of A, B and C. This prediction could be derived by a number of alternative proposals that also account for the simple case in (1). For example, it could be

maintained that the reference line is lowered by one step with each large domain boundary, or that the lowering among the clause-initial peaks is due to global declination in the sense of Pierrehumbert (1980) (see also 't Hart *et al.* 1990, Ladd 1993, Shih 2000), an effect that generally makes values later in an utterance lower than values occurring earlier, all else being equal.

Crucially, then, the different structural configuration in the *and/but* condition predicts a different relative scaling between the clauses B and C. In this structural configuration, C is not sister of B. Rather, C is sister of X = [A *and* B]. C is therefore predicted to be scaled one level lower than the reference level for X = [A *and* B]. As shown in (3b), this means that C is expected to be scaled one level lower than the initial reference level of A, but it is not predicted to be lowered relative to B.

This contrast between the two conditions was found in the English data studied by Ladd: lowering of the initial peak was observed across speakers between A and B and between B and C in the *but/and* condition. The crucial *and/but* condition also showed lowering of the initial peak between A and B across speakers and conditions; between B and C, however, there was either a smaller amount of lowering or even some raising, depending both on the speaker and on the number of accented words. (We turn to some other of Ladd's findings in §3.2.)

We consider this to be a strong argument for a direct effect of higher structure on tonal scaling. An alternative account, which postulates lowering with each large domain boundary, predicts identical lowering among clause-initial peaks in both experimental conditions. Similarly, an account in terms of global declination without sensitivity to higher structure also amounts to identical lowering. Only the hypothesis that the lowering among clause-initial peaks (partial reset) is a phonetic reflex of higher sisterhood relations correctly predicts that there is such lowering in three of the four cases of adjacent clauses, but crucially no such lowering in the fourth case, between B and C in the *and/but* condition.

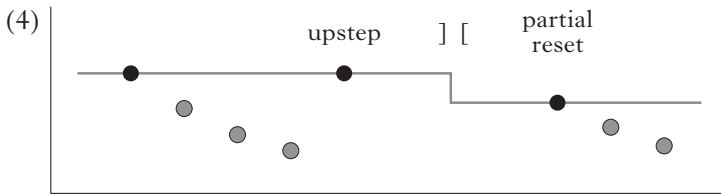
### 1.3 German intonation and upstep

We adopt the proposals for unified German ToBI transcriptions and tonal analysis in Grice & Baumann (2002) and Grice *et al.* (2005); these are further developed in Grice *et al.* (2009). Downtrends among high tones (downstep, final lowering, declination) in German have been studied experimentally by Grabe (1998), Truckenbrodt (2004, 2007b) and Féry & Kügler (2008).

Upstep in German is established in Truckenbrodt (2002, 2007b) and investigated in relation to focus in German in Féry & Kügler (2008).<sup>3</sup>

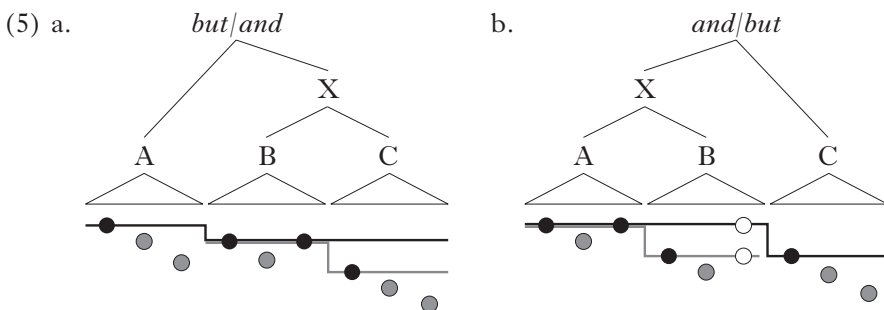
<sup>3</sup> Pierrehumbert (1980) employs the term UPSTEP in a different sense in her analysis of English, namely to denote a process of raising of the lower and upper bounds of the register by a H- phrase accent which affects a following L% or H% tone. In Grice & Baumann (2002), Grice *et al.* (2005) and Grice *et al.* (2009) the term upstep is used in a more restricted sense; there, H-<sup>^</sup>H%, where '<sup>^</sup>' stands for upstep, transcribes a sequence in which H% is higher than H-. Further phenomena related to upstep are reported in Truckenbrodt (2004, 2005).

For those speakers of German who employ the contour L\*+H (...) H% in final position in an intonational phrase (*t*) when the utterance is continued after the *t*, one or both of the two H tones are scaled to an upstepped level that is arguably similar to the initial height of that clause. Truckenbrodt's phonetic analysis of upstep is illustrated in (4). Upstep, like reset, is claimed to be a return to the phrasal reference line. Unlike reset, however, upstep occurs in domain-final position, before the phrasal reference line is downstepped. The upstepped peak is thus similar in height to the earlier clause-initial peak, and separated from the following partial reset by the lowering of the phrasal reference line between the two clauses.



Upstep on nuclear H\* accents in German is reported in Féry & Kügler (2008). There it occurs optionally on a nuclear accent in sentences with wide focus, and obligatorily on a nuclear accent on an element with narrow focus.

The present study explores for German whether upstep provides independent support for Ladd's account of the two experimental conditions. (5) illustrates the predictions for the scaling of upstep in nested structures. Upstep is expected to target the respective clause-initial height in clauses A and B in the *but/and* condition, and in clause A in the *and/but* condition. Upstep in clause B of the crucial *and/but* condition is of particular interest. The model leads to the expectation that upstep at the end of B might here target the black phrasal reference line of X=[A *and* B], just before it is lowered towards C (see the higher white circle in (5b)). However, such a scaling of upstep in B is only one predicted possibility, the other being upstep to the grey reference line of B, as shown by the lower white circle in (5b). Since the grammar allows both possibilities, we expect to find variation that is reflected in the average value of this point.



If the influence of the high reference line of X could be documented in this condition, it would provide independent evidence for Ladd's account (in the implementation we are considering here). Recall that the explanation for why C is not lowered relative to B in the *and/but* condition is that C is lowered relative to its sister X = [A *and* B]. If upstep in clause B of the *and/but* condition targets the reference line of X = [A *and* B], we would have independent evidence for the reference line in a position just before it is lowered for clause C.

## 2 The experimental methods

In the present section we illustrate the stimuli of our experiment with their prosodic and tonal analysis, and some aspects of the measurements. Additional details of the method are described in the Appendix.

### 2.1 Stimuli: AX and XC conditions

We conducted pilot studies which led to a design for our German stimuli that differs in some details from that of Ladd (1988). The two contrastive conditions of our experiments contained only two expected accents per clause. The first peak in each clause would allow for the relevant comparisons among clause-initial peaks, while the second would allow an assessment of upstep. The three clauses (A, B and C) were combined into the structures [A *während* [B *und* C]<sub>X</sub>] and [[A *und* B]<sub>X</sub> *während* C] '[A while [B and C]<sub>X</sub>]' and '[ [A and B]<sub>X</sub> while C]', comparable to Ladd's *but/and* and *and/but* conditions. The conditions are here referred to as the AX condition (A *während* X) and the XC condition (X *während* C). Each stimulus was presented in the context of a question that drew attention to the higher constituent X, which is always separated from the preceding or following clause by *während* 'while'. In the examples in (6), the clauses are separated by commas, and square brackets identify the crucial constituent X. Syllables on which accents are expected are underlined; more will be said on this below.

(6) a. *AX condition* (corresponding to Ladd's *but/and* condition):

A *während* [B *und* C]<sub>X</sub>

Warum meint Anna, dass Handwerker teurere Autos haben als Musiker?

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

Weil der Maler einen Jaguar hat, während [die Sängerin  
because the painter a Jaguar has while the singer  
einen Lada besitzt, und der Geiger einen Wartburg fährt ]  
a Lada owns and the violinist a Wartburg drives

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

- b. *XC condition* (corresponding to Ladd's *and/but* condition):

[A und B]<sub>X</sub> während C

Warum meint Anna, dass Musiker teurere Autos haben als Sportler?  
'Why does Anna think that musicians have more expensive cars than sportsmen?'

[Weil die Sängerin einen Jaguar hat, und der Geiger einen  
because the singer a Jaguar has and the violinist a

Daimler besitzt], während der Ringer einen Lada fährt  
Daimler owns while the wrestler a Lada drives

'Because the singer has a Jaguar, and the violinist owns a Daimler, while the wrestler drives a Lada.'

16 stimuli for both the AX and the XC conditions were constructed, using the words in (7).

- (7) a. *Professions*

*sportsmen*: 'Ruderer 'rower', 'Ringer 'wrestler'

*craftsmen*: 'Maler 'painter', 'Weber 'weaver'

*musicians*: 'Sängerin 'singer', 'Geiger 'violinist'

*hospital staff*: Neuro'loge 'neurologist', 'Hebamme 'midwife'

- b. *Cars*

*inexpensive*: 'Lada, 'Wartburg

*expensive*: 'Daimler, 'Jaguar

- c. *Verbs*

be'sitzt 'owns', 'fährt 'drives', 'hat 'has'

These words were permuted in the stimuli, so as to eliminate effects of individual words on F0 height (microprosodic effects or effects of the position of stress within the word) in the comparison of initial peaks in the three clauses and of final peaks in clauses A and B. In the 16 sentences of each condition, each of the eight professions occurred twice as the subject of each of the three clauses, and each of the four cars occurred four times as the object. Of the three verbs, *hat* occurred six times at the end of each of the three clauses, and the other two verbs five times.

With these stimuli, upstep on the second peak of the clauses A and B would be manifested as the absence of downstep in this position, i.e. the second peak would be similar in height to the first. To test whether the subjects would otherwise show downstep among clause-internal peaks, a further set of stimuli was included, in what we call the no-X condition. These stimuli also consisted of three clauses, but with the simpler structure [A, B *and* C]. Each clause in this condition had three expected accent locations. An example is given in (8).



(8) *No-X condition*

Warum meint Anna, dass ihre Nachbarn teure Autos haben?  
 ‘Why does Anna think that her neighbours have expensive cars?’

Weil Möller und Hummel einen Jaguar haben, Meyer und Lerner  
 einen Daimler besitzen, und Wollmann und Lehmann einen BMW  
 fahren.

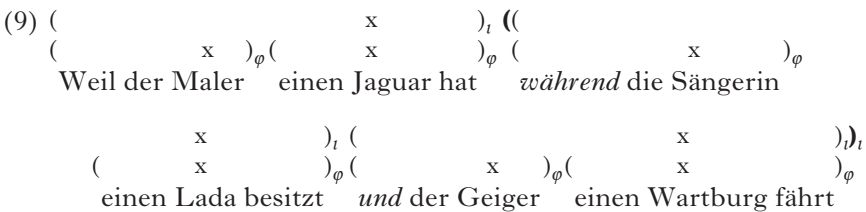
‘Because Möller and Hummel have Jaguars, Meyer and Lerner own  
 Daimlers, and Wollmann and Lehmann drive BMWs.’

All 18 stimuli of the no-X condition used the words in the example in (8). The family names, car names and verbs were again systematically permuted in order to avoid influences of individual words on the results.

**2.2 Prosodic analysis of the stimuli**

The prosodic analysis that underlines our study is illustrated in (9) with an example from the AX condition. Two phrasal prosodic levels are postulated, the lower being that of the  $\varphi$ . At this level each  $\varphi$  is headed by one accent. These domains, and the positions of the accents, can be predicted by the Sentence Accent Assignment Rule of Gussenhoven (1983, 1992), by the analysis in Truckenbrodt (2007a) or by the analysis proposed by Féry (2011), which uses recursive prosodic phrasing. The stimuli contain two  $\varphi$ ’s in each of the three clauses.

In addition, each clause is itself expected to form a higher prosodic domain. We analyse these as intonational phrases ( $\iota$ ; Beckman & Pierrehumbert 1986, Nespore & Vogel 1986, Selkirk 2005). We return to this issue in §4. The strongest (nuclear) stress in each  $\iota$  is taken to be assigned to the rightmost  $\varphi$  (Uhmann 1991, Selkirk 1995a, 2005, Truckenbrodt 2007a, Féry 2011). In a non-final  $\iota$ , the position of nuclear stress is the position of upstep in our experiment. An important issue is whether the conjunction of two clauses in the syntactic constituent X is mirrored in the prosodic structure by a further  $\iota$ . We indicate such  $\iota$ ’s in (9), and return to this question below.



The prosodic analyses of all three conditions are shown schematically in (10).



b. Warum meint Anna, dass ihre Nachbarn teure Autos haben?

‘Why does Anna think that her neighbours have expensive cars?’

L\*+H                      L\*+H                      L\*+H H%

(Weil Lehmann und Möller einen BMW haben)

L\*+H                      L\*+H                      L\*+H                      H%

(Hummel und Meyer einen Jaguar besitzen)

L\*+H                      L\*+H                      (H+)L\*                      L%

(und Lerner und Wollmann einen Daimler fahren)

‘Because Lehmann and Möller have BMWs, Hummel and Meyer own Jaguars, and Lerner and Wollman drive Daimlers.’

The analysis of the phonetic results was based on measurements of the tones shown in (11): non-final L\*+H pitch accents and L\* or H+L\* pitch accents, as well as H% boundary tones at the end of non-final clauses, and a L% boundary tone in utterance-final position.

The phonetic analysis included only those recordings that showed clause-initial L\*+H accents in all three clauses, as well as H% boundary tones at the end of both the first and second clauses. When the second pitch accent of clauses A and B was not a L\*+H rise, it was not included in the analysis, even though the remainder of the token was.

In most cases, our criteria allowed at least 14 of the 16 recordings for both the AX and the XC conditions to be included in the measurements of a given speaker. The two exceptions were the XC condition for speaker S2, where only six of the 16 tokens could be measured, and the AX condition for speaker S3 (seven tokens). In clauses A and B of the AX and XC conditions taken together, one measurement of the nuclear accent (L\*+H) had to be skipped. See Appendix: §4 for further details of the data retained for analysis.

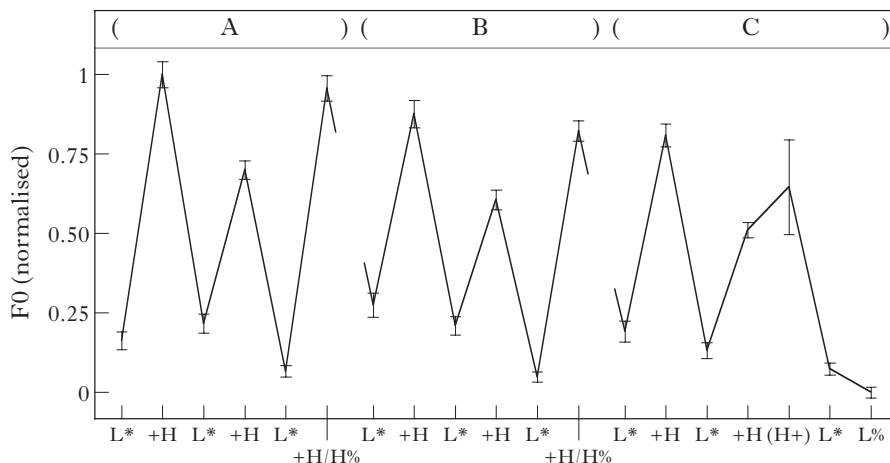
For the measurements of upstep in the sequence L\*+H H%, only one phonetic value, the highest in the area of +H H%, enters into the phonetic evaluations below. As shown in Truckenbrodt (2007b), upstep in such tonal sequences may occur either on the clause-final H% or on the preceding +H of the nuclear rise, or on both. Therefore, the best approximation to the upstepped height seems to be the highest value in this area for each token.

### 3 Results and discussion: effects of hierarchical structure on phonetic height

In this section we show the results from the no-X condition (§3.1), our crucial replications of Ladd’s (1988) findings with the results from the AX and XC conditions (§3.2), and our findings for upstep in the AX and XC conditions (§3.3).

The measured values were normalised, using the linear transformation in (12).

$$(12) \text{transformed\_value} = (\text{measured\_value} - L\%[\text{speaker}]) / (\text{H1}[\text{speaker}] - L\%[\text{speaker}])$$



*Figure 1*

Pooled normalised measurements and 95% confidence intervals of the no-X condition. The brackets at the top of the plot separate the values within the three clauses A, B and C. '(H+)' is a measurement point related to a final H+L\* in clause C only for S1, S3 and S4, who employ a final H+L\* in clause C.

For the AX and XC conditions, L%[speaker] and H1[speaker] were calculated on the pooled values of these two conditions. L%[speaker] is the average of the L% value (final in clause C) for a speaker, and H1[speaker] the average of the H1 values (initial in clause A).

The crucial results rest on the comparison of the AX and XC conditions. The no-X condition is longer than the other two conditions. The normalisation, including the values of L%[speaker] and H1[speaker], was therefore calculated separately for the no-X condition.

### 3.1 No-X condition

Figure 1 shows the averaged measurements of the recordings of the no-X condition, normalised and pooled across the five speakers. The no-X condition allows us to make the following points. First, there is a clear pattern of downstep between the first and second H peaks in each clause. Second, the high values at the end of the clauses A and B do not continue the downstepping pattern of the first two peaks of the clause. Rather, the last high values return to approximately the height of the initial peak of their clauses, i.e. they display upstep. Together, these two observations show that there is *t*-internal downstep and *t*-final upstep for the speakers we recorded.

The AX and XC conditions contain only an initial and a final peak in each clause. In these cases, if the second, clause-final peak and the first, clause-initial peak are similar in height, this might simply seem like the

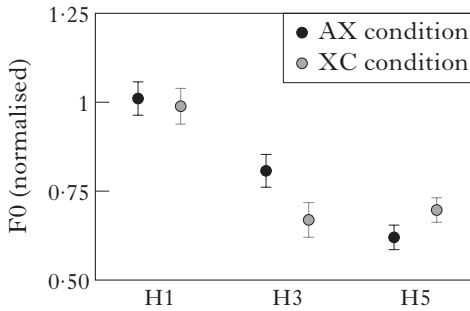


Figure 2

Averaged values and 95% confidence intervals for the clause-initial peaks in the normalised pooled values of speakers S1–S5. H1, H3 and H5 are the respective initial peaks of the clauses A, B and C.

absence of downstep or upstep. However, the no-X condition shows that the recorded speakers show a pattern of downstep and upstep relative to  $\iota$ , as has been described in previous studies (Truckenbrodt 2007b, Féry & Kügler 2008). If, therefore, the heights of the two peaks in a non-final clause in the AX and XC conditions are similar, we are justified in analysing this as upstep preceding an  $\iota$  boundary.

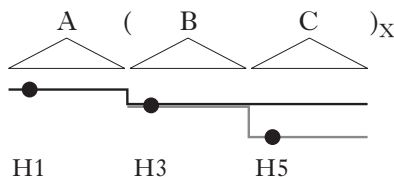
We have not analysed the relation between the three clauses in the no-X condition. We think that it is possible that individual speakers superimpose their own structure [AB]C or A[BC], since none of these options is enforced by the experiment (see Kentner & Féry 2013 for evidence that speakers do impose such a structure on a ternary sequence of names).

### 3.2 The relative height of the clause-initial peaks: the effect of hierarchical structure on tonal scaling

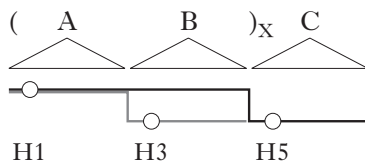
Figure 2 shows the comparison of the three clause-initial peaks for the pooled normalised data from all five speakers. These clause-initial peaks allow for an assessment of the scaling of the three clauses relative to each other. The figure shows a replication of the crucial result of Ladd (1988) reviewed above. In the AX condition (Ladd's *but/and* condition), there is lowering not just between the initial peaks of clauses A and B (H1 and H3 respectively) but also between those of clauses B and C (H3 and H5). By contrast, the XC condition (Ladd's *and/but* condition) shows lowering between A and B, but not between B and C. This is confirmation both of Ladd's claim that higher structure affects tonal scaling and of the particular effect of higher structure on tonal scaling postulated by Ladd: lowering among large domains affects nodes that are structural sisters in the representation. As shown in (13b), this correctly predicts the absence of lowering between clauses B and C in the XC condition, where C is not downstepped relative to B, but rather to its structural sister  $X = [A \text{ and } B]$  and thus relative to the initial height of A. At the

same time, it correctly predicts lowering among all other adjacent clauses, including lowering between B and C in the AX condition in (13a), where C is a structural sister of B.

(13) a. *AX condition*



b. *XC condition*



Other results from Ladd (1988) were also replicated. These are also discussed with reference to Fig. 2. For one thing, as in Ladd's data, the utterance-initial peaks (H1) in both conditions are similar in height. Ladd's model does not lead us to expect a difference in utterance-initial height, and it is thus reasonable to expect that a similar utterance-initial level is used by the speakers in both conditions.

Second, there is a clear difference in the height of the initial peak in clause B in the two experimental conditions. In Fig. 2, this is visible at point H3, which is higher in the AX condition than in the XC condition. The difference is highly significant.<sup>4</sup> The difference was also found by Ladd, but is not predicted by his model, as we have seen, since clause B is lowered by one step relative to clause A in both conditions. Ladd (1988: 541) describes this in a formulation that correlates boundary size with the strength of the reset: 'clause-initial accent peaks are higher following a stronger boundary'.<sup>5</sup> Here we relate this to final lowering, which is known to be a factor in sequences of downstepping accents. As far as their non-final accents are concerned, these sequences have the shape of exponential decay towards an abstract reference line; however, the final accent deviates from this pattern and shows a downstep that is deeper than expected (Lieberman & Pierrehumbert 1984). The relevant sense of 'final' could be utterance- or phrase-final or 'final in the downstep sequence' in the English data of Lieberman & Pierrehumbert (1984) and in the Mexican Spanish data of Prieto *et al.* (1996). Truckenbrodt (2004) argues that at least some amount of final lowering is found in accent sequences in German in the position 'final in the downstep sequence', even where this is not final in an utterance or final in an *t*. In the cases at hand this account does not make any additional predictions.

<sup>4</sup> H3 in AX = 0.80, H3 in XC = 0.69;  $t_{66} = 4.14$ ,  $p < 0.001$  in a two-tail *t*-test. Since the total number of normalised measurements in position H3 is 67 in AX and 66 in XC, the average of XC was added to the XC values in this calculation.

<sup>5</sup> In Féry & Truckenbrodt (2005), the difference in the height of H3 in the two conditions is accounted for by a principle that we call 'the deeper the steeper', which specifies that the more deeply embedded the constituents are, the steeper the downstep is.

	AX condition			XC condition		
	[ ] <sub>X</sub>			[ ] <sub>X</sub>		
	H1	H3	H5	H1	H3	H5
final lowering			final in X final in utterance		final in X	final in utterance
downstep		one step down	two steps down		one step down	one step down

Table I

Predictions of final lowering and downstep at the phrasal level.

We suggest generalising the concept of final lowering from the height of accents to the height of phrasal reference lines. Table I provides an overview of different domain-final positions in our data. As shown, H3 (i.e. the reference height of clause and intonational phrase B) is final in X in the XC condition, but not in the AX condition. If the reference line of B undergoes final lowering in the XC condition because it is final within the large X, then final lowering can account for the difference between the two values of H3 in Fig. 2.

A similar account can be given for the distinction in height between H3 of the AX condition and H5 of the XC condition in Fig. 2. These each undergo only one step of downstep among phrases (in the case of H5 of the XC condition, this is the crucial step of lowering relative to the constituent X = [AB]). Therefore downstep among phrases alone would lead us to expect that they would be of similar height. The fact that H5 of the XC condition is nevertheless lower than H3 of the AX condition can be explained in terms of final lowering. Clause B of the AX condition, as we have seen, is not final in any sense. Clause C of the XC condition, on the other hand, is final in the utterance. Final lowering in its application to phrases can thus explain that H5 in the XC condition is lower than H3 in the AX condition. Notice that (3), (5) and (13) display not only downstep among phrases but also final lowering applied to phrases.

Taken together, downstep among phrases and final lowering applied to phrases, both crucially affected by the constituent X, provide a good analysis of the values in Fig. 2. The analysis thus strengthens the claim of the relevance of higher structure for tonal scaling in terms both of the effect of X on downstep among phrases and of the effect of X on final lowering applied to phrases.

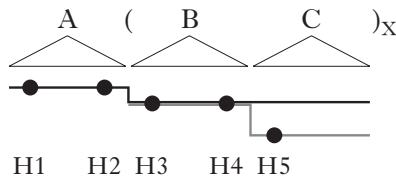
In summary, we have replicated the core findings of Ladd (1988) in regard to clause-initial peaks in the two experimental conditions A[BC] and [AB]C. The most important aspect is that C is lowered relative to B in A[BC], where C is sister to B, while C is not lowered relative to B in

[AB]C, where C is not sister to B. In the latter case, C is instead lowered relative to its sister, [AB]. This provides confirmation of a core aspect of Ladd's proposal about the effect of hierarchical structure on tonal scaling: lowering among higher domains is systematic, and tied to a structural configuration of sisterhood. We have also replicated other aspects of his results, which we analyse in terms of final lowering applied to phrases. We have argued that this provides additional support for the role of the higher structure in tonal scaling.

### 3.3 The relative height of the upstepped values

(14) shows our expectations for the upstepped values, and illustrates our test question. We expect the upstepped values to be of the same height as the immediately preceding clause-initial values in clauses A and B of the AX condition ( $|H1| = |H2|$ ,  $|H3| = |H4|$ ), and in clause A of the XC condition ( $|H1| = |H2|$ ). To the extent that this is the case, we can test the question we are interested in, which concerns upstep in clause B of the XC condition. The relevant tone is the upstepped H4 of the XC condition. Does this upstepped tone target the reference line at the initial height of B ( $|H3| = |H4|$  in XC), or does it target the higher abstract reference line of the constituent X, just before this line is lowered for clause C in the XC condition ( $|H1| = |H4|$  in XC)?

(14) a. *AX condition*



b. *XC condition*

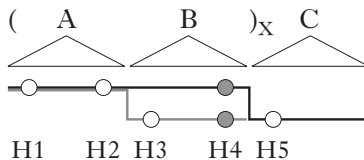


Table II gives a comparison for each speaker of the clause-initial peaks H1 and H3 with the upstepped values H2 and H4. In the first three columns we expect no difference in height between a clause-initial peak and the following upstepped peak. The last two columns, which compare H4 with both H3 and H1 in the XC condition, bear on the test case.

Notice that speaker S5 shows significant differences in the first three columns. For this speaker, then, upstep does not normally target the clause-initial height in this data; rather, the upstepped value is unexpectedly high. We therefore put this speaker aside for the moment, and return to him below. For the remaining four speakers, there are no significant differences in the first three columns of Table I after Bonferroni adjustment, i.e. upstep is broadly similar in height to the immediately preceding clause-initial value in the three cases where we expect height to be the same. Figure 3 shows pooled normalised values of clause-initial peaks and upstepped peaks for these four speakers, S1–S4. The plots of



	AX condition				XC condition					
	H1	H2	H3	H4	H1	H2	H3	H4	H1	H4
S1	282	281	245	244	278	279	249	258	278	258
	$t_{13} = 0.17$		$t_{13} = 0.46$		$t_{15} = -0.21$		$t_{15} = -2.08$		$t_{15} = 5.47$	
	$p > 0.05$		$p > 0.05$		$p > 0.05$		$p > 0.05$		$p < 0.001^*$	
S2	261	262	244	233	256	260	236	237	256	237
	$t_{14} = -0.21$		$t_{14} = 2.88$		$t_5 = -0.45$		$t_5 = -0.18$		$t_5 = 4.13$	
	$p > 0.05$		$p = 0.012$		$p > 0.05$		$p > 0.05$		$p = 0.009^*$	
S3	273	269	234	262	262	268	204	260	262	260
	$t_6 = 0.59$		$t_6 = -2.90$		$t_{12} = -0.68$		$t_{12} = -7.53$		$t_{12} = 0.17$	
	$p > 0.05$		$p = 0.027$		$p > 0.05$		$p < 0.001^*$		$p > 0.05$	
S4	140	146	134	137	141	146	121	139	141	139
	$t_{14} = -1.68$		$t_{14} = -1.14$		$t_{14} = -1.26$		$t_{14} = -4.99$		$t_{14} = 0.61$	
	$p > 0.05$		$p > 0.05$		$p > 0.05$		$p < 0.001^*$		$p > 0.05$	
S5	163	183	156	169	163	186	151	171	163	171
	$t_{15} = -7.43$		$t_{15} = -6.28$		$t_{15} = -6.87$		$t_{15} = -8.60$		$t_{15} = -2.76$	
	$p < 0.001^*$		$p < 0.001^*$		$p < 0.001^*$		$p < 0.001^*$		$p = 0.014$	

Table II

Comparisons (in Hz) of upstepped values H2 and H4 with earlier clause-initial values H1 and H3, using paired-sample *t*-tests.

Asterisks highlight differences that are significant after Bonferroni adjustment for each speaker ( $p < 0.05/5$ , i.e.  $p < 0.01$ ).

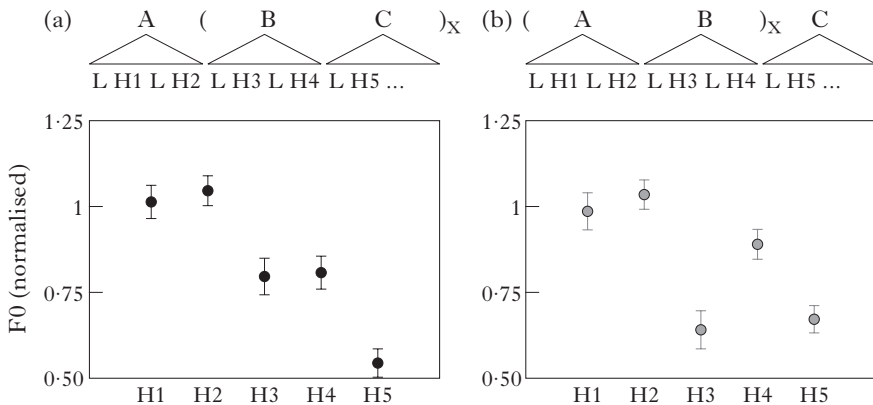


Figure 3

95% confidence intervals for upstepped values H2 (clause A) and H4 (clause B) in relation to the three clause-initial values H1, H3 and H5 in the normalised pooled data of speakers S1–S4. (a) AX condition ( $n = 51$ ); (b) XC condition ( $n = 50$ ).

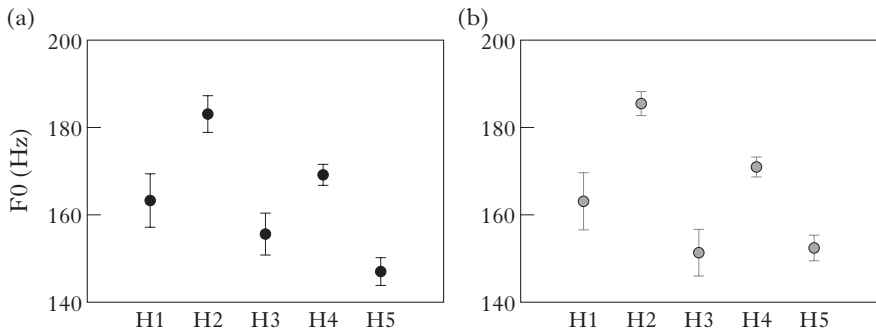


Figure 4

Averaged values and 95% confidence intervals for the clause-initial and upstepped peaks of speaker S5. (a) AX condition ( $n = 16$ ); (b) XC condition ( $n = 16$ ).

Fig. 3 show that  $|H1| = |H2|$  and  $|H3| = |H4|$  in the AX condition, and  $|H1| = |H2|$  in the XC condition. For the crucial test case, it can be seen that H4 in the XC condition is much higher than the preceding H3. This strongly suggests that the high reference line of X in the XC condition plays a role as a target for upstep in H4.

H4 is also lower than H1 in the XC condition in Fig. 3. The two right-most columns of Table I suggest that speakers show individual preferences as to the scaling of H4 in the XC condition. These results are compatible with speakers S1 and S2 scaling H4 to the lower reference line of B in (14b), and speakers S3 and S4 scaling H4 to the higher reference line of [AB]. S3 and S4 are thus primarily responsible for the greater height of H4 relative to H3 in Fig. 3b, the XC condition, and S1 and S2 are primarily responsible for the lower height of H4 relative to H1 in the same plot. This is compatible with the model we are pursuing, which allows both scaling options for H4 in (14).

For completeness, we return briefly to speaker S5, whose data is given in Fig. 4. It is clear that in both conditions each upstepped value H2 and H4 is higher than the peak immediately preceding it. We interpret this in terms of the suggestion of Kentner & Féry (2013), who formulate a principle called Proximity to express the fact that speakers reduce the first boundary in a phrase grouping two constituents in the same  $\phi$ . Anti-proximity accounts for the fact that speakers increase a boundary before a following boundary to express the separation of constituents. Petrone *et al.* (2014) argue for an understanding of Proximity as a strategy that speakers may use over and above the default prosodic mapping: where Y and Z are otherwise mapped to separate  $\phi$ 's, a constituent [YZ] can be made salient by being mapped to a single  $\phi$ . In the current experiment, speaker S5 seems to have made such an effort in the domains [H1 H2] and [H3 H4].

## 4 On the prosodic representation of the two experimental conditions

In this section we discuss the recursive *t* phrasing of the clauses A, B, C and X in the experiment.

Ladd (1986, 2008: 296ff) proposes that the constituent that is here labelled X forms a compound *t*, consisting of two embedded *t*'s. This is a compound *t* formed by B and C in the AX condition, A[BC], and a compound *t* formed by A and B in the XC condition, [AB]C. We think that this is correct for principled reasons. Much of the work in prosodic phonology since Selkirk (1980) and Nespor & Vogel (1986) (see Truckenbrodt 2007a and Selkirk 2011 for review) has been guided by the Indirect Reference Hypothesis (Inkelas 1989: 9), according to which phonological rules may not access syntactic structure directly. Rather, syntactic structure, in these accounts, plays a role in conditioning a prosodic constituent structure, and only this prosodic constituent structure may be referred to by the phonology. On standard assumptions, the interface between phonology and phonetics occurs after the mapping from syntax to phonology. Therefore, the phonetic implementation (such as the determination of tonal height) will not have access to the syntactic structure either. Syntactic structure can only indirectly affect phonetic scaling, to the extent that it leads to prosodic structure that affects the phonetic implementation. (See Kabagema-Bilan *et al.* 2011 and Katz & Selkirk 2011 for discussion.) It also seems that this position converges with the standard assumption that it is prosodic structure that affects phonetic implementation (Kent & Netsell 1971, Pierrehumbert 1980, Pierrehumbert & Beckman 1988, Fougeron & Keating 1997 and much recent work).

On this view, Ladd's (1988) results and ours provide empirical support for recursive *t* phrasing: (A)<sub>*t*</sub>((B)<sub>*t*</sub>(C)<sub>*t*</sub>)<sub>*t*</sub> in the AX condition and ((A)<sub>*t*</sub>(B)<sub>*t*</sub>)<sub>*t*</sub>(C)<sub>*t*</sub> in the XC condition. There is evidence that each of the individual clauses A, B and C constitute a separate *t* from the observation that they are domains for partial F0 resetting in Ladd's experiment, and domains for upstep in our experiment. At the same time, X, comprising B and C in the AX condition and A and B in the XC condition, must also be mapped to larger and complex *t*'s, so that the distinctions between the two experimental conditions can be accounted for.

How, then, does the syntax–phonology mapping derive the claim that each of the syntactic constituents X, A, B and C turn into coextensive *t*'s? Since this is a mapping to an isomorphic structure, one might think that it should be straightforward. However, it turns out that not all mapping accounts are equipped to derive the isomorphic recursive structure.

The three clauses in Ladd's experiment are root sentences in the sense of Downing (1970): they are not embedded in a higher clause that has a predicate of its own. Similarly, the combined clauses [A and B] in the *and/but* condition and [B and C] in the *but/and* condition are root sentences, as is the combination of all three clauses in both conditions. We think that

our stimuli are best viewed in the same way. Although our sentences have the typical shape of embedded German clauses (verb-final word order), they occur without overt embedding. Downing (1970) argues that root sentences are obligatorily separated by  $\iota$  boundaries from surrounding material at their left and right edges. If, however, we think of this suggestion in terms of syntax–prosody alignment (Selkirk 1986, 1995b), the two experimental conditions are wrongly mapped to identical prosodic structures, as shown in (15). (15a) is a sketch of the syntactic structure, while (15b) shows the phrasing derived by left and right alignment of clause boundaries with  $\iota$  boundaries.

(15) *Insufficient mapping by syntax–prosody alignment of clauses with  $\iota$  boundaries*

<i>AX condition</i>	<i>XC condition</i>
a. A [B C] <sub>X</sub>	[A B] <sub>X</sub> C
b. ( ) <sub><math>\iota</math></sub> ( ) <sub><math>\iota</math></sub> ( ) <sub><math>\iota</math></sub>	( ) <sub><math>\iota</math></sub> ( ) <sub><math>\iota</math></sub> ( ) <sub><math>\iota</math></sub>

Each of the clauses A, B and C form an  $\iota$ , whose boundaries also satisfy the alignment requirement for the higher constituents. In particular, X = [BC] in the AX condition now also has  $\iota$  boundaries at both edges, as does the constituent X = [AB] in the XC condition. The account would not derive the prosodic distinction between the two conditions that gives rise to the different clause-initial height relations in Ladd’s experiment or in ours.

No improvement can be made by the addition of a WRAP constraint to the account. The constraint WRAPXP interacts with edge-alignment of XPs in Truckenbrodt (1995, 1999), and a WRAP constraint for the relation between clauses and  $\iota$ ’s is formulated in Selkirk (2005) and Truckenbrodt (2005), with differences in detail that are not relevant here. The requirement, applied to the structures under consideration, is in both formulations that each clause must be contained in an  $\iota$ . The addition of this constraint would derive the prosodic structures in (16b) if wrapping suppresses alignment, and the prosodic structure in (16c) if wrapping does not suppress alignment.

(16) *Wrong mapping by alignment and wrapping*

<i>AX condition</i>	<i>XC condition</i>
a. A [B C] <sub>X</sub>	[A B] <sub>X</sub> C
b. ( ) <sub><math>\iota</math></sub>	( ) <sub><math>\iota</math></sub>
c. (( ) <sub><math>\iota</math></sub> ) <sub><math>\iota</math></sub> ( ) <sub><math>\iota</math></sub> ( ) <sub><math>\iota</math></sub>	(( ) <sub><math>\iota</math></sub> ) <sub><math>\iota</math></sub> ( ) <sub><math>\iota</math></sub> ( ) <sub><math>\iota</math></sub>

In both structures the single overarching  $\iota$  would satisfy the wrap requirement for all lower clauses in both conditions. In particular, the constituent X = [BC] in the AX condition and the constituent X = [AB] in the XC condition are correctly wrapped by the larger  $\iota$  in the structures in (16).

Wrapping, like alignment, does not provide an incentive to map these intermediate constituents to the prosodic structure. However, these are the constituents that crucially distinguish the two conditions.

An account that correctly distinguishes the two structures is the Match theory of Selkirk (2011). This theory postulates that syntactic words are mapped to identical prosodic words, syntactic XPs are mapped to identical  $\phi$ 's and syntactic clauses are mapped to identical  $\iota$ 's. In addition, Selkirk suggests replacing Downing's (1970) notion of root sentence with that of illocutionary clause, following up on remarks in Potts (2005) about the connection with speech acts. Her theory includes an additional constraint that maps illocutionary clauses to identical  $\iota$ 's. In the context of our paper, it is not important whether all clauses or only illocutionary clauses are matched to  $\iota$ 's. For concreteness, we employ the constraint that matches illocutionary clauses to  $\iota$ 's, given in (17).

(17) MATCHILLOCUTIONARYCLAUSE (Selkirk 2011)

The left and right edges of an illocutionary clause must correspond to the left and right edges of an  $\iota$ .

In this account, each illocutionary clause (or root sentence) is directly mapped to an  $\iota$ . This has the desired effect, that the constituents  $X = [BC]$  in the AX condition and  $X = [AB]$  in the XC condition are not only syntactic constituents, but also matching prosodic constituents, as shown in (18).

(18) *Correct mapping by matching clauses and  $\iota$  boundaries*

<i>AX condition</i>	<i>XC condition</i>
a. A [B C] <sub>X</sub>	[A B] <sub>X</sub> C
b. (( ) <sub><math>\iota</math></sub> (( ) <sub><math>\iota</math></sub> ( ) <sub><math>\iota</math></sub> ) <sub><math>\iota</math></sub>	(( ( ) <sub><math>\iota</math></sub> ( ) <sub><math>\iota</math></sub> ) <sub><math>\iota</math></sub> ( ) <sub><math>\iota</math></sub> ) <sub><math>\iota</math></sub>

We therefore think that Ladd's results and ours provide empirical support for the move from Downing's formulations in terms of edges alone to the stronger requirement of Match theory that the two edges that are aligned must be edges of the same prosodic constituent, the one that is matched to the relevant illocutionary clause (or root sentence).

Selkirk's Match theory assumes a good deal of recursion in prosodic structure, following an early suggestion of Ladd (1986). In other earlier prosodic accounts, recursion was either assumed not to exist (Nespor & Vogel 1986, Selkirk 1986), or was derived as an exception (Selkirk 1995b).

An exception is the model proposed by Féry (2010, 2011, 2015) and Kentner & Féry (2013), which rests on an analysis in which prosodic structure is recursive, not only at  $\phi$  level but also at  $\iota$  level. In Féry (2015), it is shown in a Match model that embedded relative clauses and complement clauses often avoid recursivity by extraposition, but that recursive  $\iota$ 's are

allowed in the prosody of German. Féry (2010) and Kentner & Féry (2013) show that syntactic and prosodic recursivity mirror each other.<sup>6</sup>

We believe that our conclusion holds even in the presence of a functional motivation for indicating boundaries. Our experimental set-up was designed to draw attention to the semantic relevance of the X constituent in both conditions. We think, however, that this can only lead to an *ι* corresponding to X if the grammar contains an incentive to map X to an *ι* to begin with. Without such a mapping relation in the grammar, there is no sense in which the phonetic cues for the *ι* for X = [AB] provide a cue for a corresponding syntactic structure to the listener.

In conclusion, each root sentence (or illocutionary clause or speech act) is mapped to an *ι*. This is a strengthened requirement of the suggestion of Downing (1970) that root sentences have *ι* boundaries at their left and right edges. The way in which Downing's suggestion needs to be strengthened supports the Match format of relating syntactic and prosodic constituents in Selkirk (2011).

## 5 Conclusion

In our experiment, we have replicated the core findings of Ladd (1988) in regard to the scaling of clause-initial peaks in nested structures. In addition, we have shown that structural distinctions of the kind investigated by Ladd are also involved in the scaling of German clause-final upstep, in a way that confirms Ladd's account. Our findings lend support to the following conclusions.

(i) Downstep applies (in the typical case) among hierarchical sister nodes, allowing for the application of downstep within downstep, as postulated by Ladd (1988, 1990).

(ii) These downstep relations can be sensibly modelled using the phrasal reference lines of van den Berg *et al.* (1992) (see also the extension of this model in Truckenbrodt 2007b), with reference lines which are constant for a given higher constituent, and lowered among sister nodes.

(iii) Clause-final upstep in German, where it applies, involves scaling on this phrasal reference line, as suggested in Truckenbrodt (2002, 2007b).

(iv) The data investigated here require an isomorphic mapping from root sentences (or illocutionary clauses, or speech acts) to *ι*'s. The mapping needs to distinguish A[BC]<sub>X</sub> from [AB]<sub>X</sub>C, where A, B, C and X are root sentences. To achieve this, the account of Downing (1970), in which root sentences are bounded by *ι* edges, needs to be strengthened. The right kind of strengthening cannot be achieved in terms of alignment in interaction with wrapping, but can be within the Match theory of Selkirk (2011) and a recursive account of prosodic structure (Féry 2010, 2011, 2015): (illocutionary) clauses are matched to *ι*'s.

<sup>6</sup> The relation between speech acts and *ι*'s put forward in Selkirk (2011) is explored in detail in German in Truckenbrodt (2015), which provides support for Selkirk's MATCH constraint.

## Appendix: Details of the method

### 1 Participants

The five speakers were students at the University of Potsdam, all in their twenties and from the northern half of Germany. They were monolingual native speakers of German; the first time they had learned a second language was in school. S1–S3 were female, S4–S5 were male. They were reimbursed for their participation.

### 2 Recordings

The recordings were made in a quiet room on a DAT tape recorder. The experimenter left the participant alone in the room with the tape recorder running, after brief initial instructions as to how to begin and end the session. The participants went through the experiment in the form of a PowerPoint presentation in a self-paced manner. The instructions familiarised the participants with the procedure, and made them practise the procedure with four examples. The procedure for each stimulus consisted of two steps. The question and the answer were presented together on the screen, and the participant was asked to take the time to understand the relation in meaning between the question and the answer. When the participant pressed the 'forward' key of the presentation, the visual display remained unchanged, and a prerecorded voice read the question. The participant then read aloud the stimulus as an answer to that question. The set-up and the instructions included the option of repeating the two steps for a particular stimulus in case of hesitation during production, or if the production was not felt to be natural. The instructions further specified that the stimulus should be produced in a normal, narrative tone of voice, and at a normal rate of speech (*zügige, nicht verweilende Sprechgeschwindigkeit*). They also indicated that there were two parts of the experiment, gave the options of re-reading the instructions and of going over the practice recordings again, and made suggestions about taking breaks.

The stimuli from the AX and XC conditions were presented first, pseudo-randomised as a group. No fillers were employed in this group. For one thing, the task of concentrating on the connection between question and answer before each utterance required close attention and pausing; it seemed to us that this would itself prevent repetitive outputs from arising. For another, this concentration is demanding; if fillers of the same kind had been added, the participants might not have been able to maintain concentration for all the stimuli.

After an optional break, participants were presented with the stimuli from the no-X condition. These stimuli all employed the same prosodic pattern, and there was no special cognitive task preceding the production of each stimulus. Filler sentences were therefore interspersed, to minimise the occurrence of repetitive routines.

### 3 Measurements

The recordings were analysed using Praat. The recordings were manually divided into labelled substrings with the help of spectrograms. This was performed by student research assistants at the University of Potsdam, and checked by the authors. The divisions assigned included accent-domain

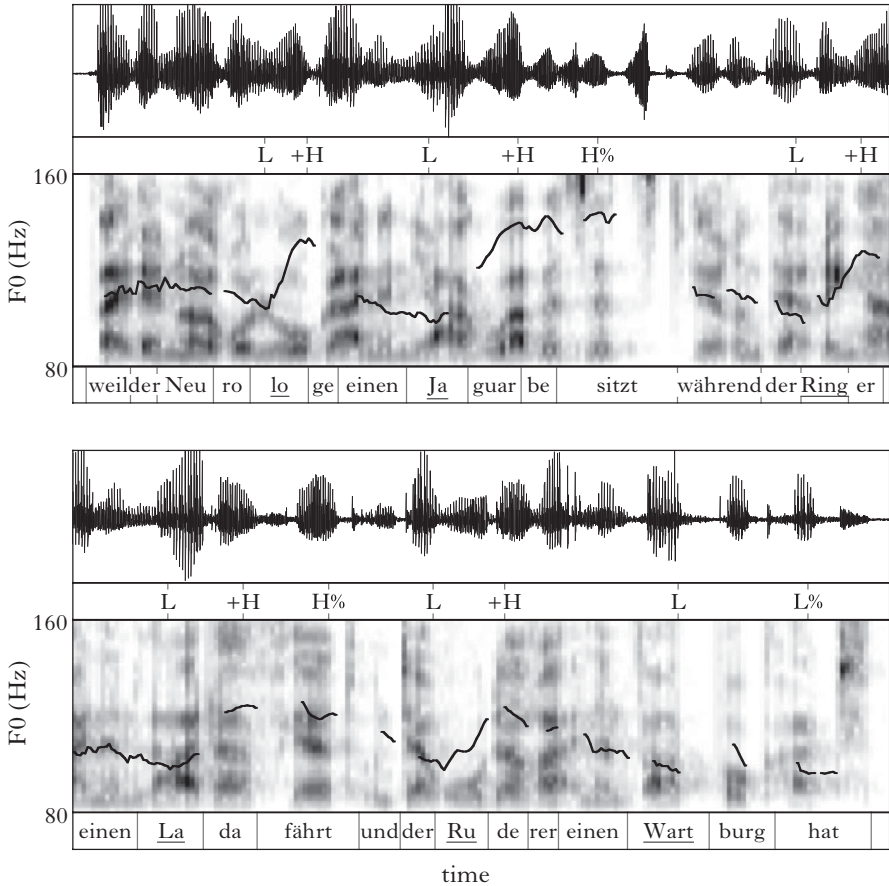


Figure 5

Pitch track from the AX condition for *weil der Neurologe einen Jaguar besitzt, während der Ringer einen Lada fährt, und der Ruderer einen Wartburg hat* 'Because the neurologist owns a Jaguar while the wrestler drives a Lada and the rower has a Wartburg' (Speaker S4). Accented syllables are underlined.

boundaries (Gussenhoven 1983, Ladd 1983b) as well as beginnings and ends of the accented syllables. Relevant possible accents were H\*, L\*, L\*+H, L+H\* H+L\* and H+!H\*. These were assigned jointly by the two authors, based on a combination of auditory impression and the F0 contour. There were no cases of disagreement.

Acoustically, a postulated L\*+H accent in clause-initial position showed an F0 minimum around the onset of the stressed syllable (this F0 minimum was analysed as the L\* part), and a following rise terminating before the end of the  $\varphi$  (the F0 maximum of the remainder of the  $\varphi$  was analysed as the +H part). A postulated L\*+H accent in clause-final utterance-medial position showed a similar position for the F0 minimum, a following rise and a high turning



point at which the rise turned into a horizontal (or more gradually rising) plateau (this further turning point was then analysed as the +H part of the pitch accent). A final H% edge tone was assigned at the end of such a final plateau, on the clause-final verb. Accents perceived as clearly high or falling on the stressed syllable were not included in the category L\*+H. In the few cases where it was not clear perceptually whether there was low or high tone on the stressed syllable, a L\*+H was assigned, since this was not in conflict with the auditory impression. In cases of distortions of the pitch track due to the presence of an obstruent, the measurement was made in an area that was judged to be just outside the distortion caused by the obstruent.

#### 4 Exclusion of values

As explained in §2.3, the criterion for including a recording in the evaluation of phonetic scaling was the assignment of clause-initial L\*+H pitch accents with the L\* part just before or early in the stressed syllable in all three clauses, as well as H% at the end of both the first and second clauses. An example of a pitch track in which this criterion is satisfied is given in Fig. 5. Here the non-final pitch accents are labelled L+H, and the final pitch accent is L.

The criteria for inclusion were applied jointly by the two authors. The second column in Table III shows that most recordings of the no-X condition met these fairly strict criteria. The third column shows that L\*+H was also assigned with great regularity in the nuclear position of clauses A and B (otherwise, individual measurements were skipped).

The same criteria for inclusion in the phonetic analysis were used in the AX and XC conditions. The frequency of occurrence of the tones is shown in Table IV.

	three clause-initial L*+H and two clause-final H% (n = 18)	L*+H on nuclear accents of clauses A, B (n = 36)	utterance-final pitch accents
S1	16 (2 different accents)	36	16 H+L*, 2 H+!H*
S2	18	36	11 L*, 7 L*+H
S3	16 (1 different accent; 1 without H%)	35	9 H+L*, 7 H+!H*, 2 H*
S4	18	34	10 H+L*, 6 L*, 2 H+!H*
S5	18	36	17 L*, 1 H*

*Table III*

Frequency of the core characteristics of the no X condition. The second column shows the number of recordings that met the criteria for inclusion in the phonetic analysis. Reasons for exclusion are given in parentheses. The third column shows the assignment of L\*+H in the second and third positions of clauses A and B. The fourth column shows the utterance-final pitch accents that were assigned.

	condition	three clause-initial L*+H and two clause-final H% ( <i>n</i> = 16)	L*+H in upstep position in clauses A, B ( <i>n</i> = 32)	utterance-final pitch accents
S1	AX	14 (2 without H%)	32	15 H+L*, 1 H+!H*
	XC	16	32	14 H+L*, 2 H+!H*
S2	AX	15 (1 without H%)	32	11 L*, 5 L*+H
	XC	6 (10 without H% in B)	32	13 L*, 3 L*+H
S3	AX	7 (9 different accents)	32	13 H+L*, 3 H+!H*
	XC	13 (2 different accents; 1 without H%)	32	12 H+L*, 4 H+!H*
S4	AX	15 (1 without H%)	31	13 H+L*, 3 L*
	XC	15 (1 without H%)	32	11 H+L*, 5 L*
S5	AX	16	32	16 L*
	XC	16	32	16 L*

*Table IV*

Frequency of the core characteristics of the AX and XC conditions. The numbers of recordings that met the criteria for inclusion in the phonetic analysis are shown in the third column. The fourth column shows the assignment of L\*+H in the second positions of clauses A and B. The fifth column shows the utterance-final pitch accents that were assigned.

As mentioned in §2.3, there were two cases in which fewer than half of the recordings could be included in the phonetic analysis. Speaker S2 often marked the larger internal boundary of the XC condition with a L% preceded by an accentual fall instead of a rise followed by H%, and S3 sometimes used a different pitch accent in clause-initial positions (H\* was assigned in most of these cases).

## REFERENCES

- Bartels, Christine (1999). *The intonation of English statements and questions: a compositional interpretation*. New York & London: Garland.
- Beckman, Mary E. & Janet B. Pierrehumbert (1986). Intonational structure in Japanese and English. *Phonology Yearbook* 3. 255–309.
- Berg, Rob van den, Carlos Gussenhoven & Toni Rietveld (1992). Downstep in Dutch: implications for a model. In Gerard J. Docherty & D. Robert Ladd (eds.) *Papers in laboratory phonology II: gesture, segment, prosody*. Cambridge: Cambridge University Press. 335–359.
- Bruce, Gösta (1977). *Swedish word accents in sentence perspective*. Lund: Gleerup.
- Downing, Bruce T. (1970). *Syntactic structure and phonological phrasing in English*. PhD dissertation, University of Texas.
- Féry, Caroline (2010). Syntax, information structure, embedded prosodic phrasing, and the relational scaling of pitch accents. In Nomi Erteschik-Shir & Lisa Rochman (eds.) *The sound patterns of syntax*. Oxford: Oxford University Press. 271–290.

- Féry, Caroline (2011). German sentence accents and embedded prosodic phrases. *Lingua* **121**. 1906–1922.
- Féry, Caroline (2015). Extraposition and prosodic monsters in German. In Lyn Frazier & Ted Gibson (eds.) *Explicit and implicit prosody in sentence processing*. New York: Springer.
- Féry, Caroline & Frank Kügler (2008). Pitch accent scaling on given, new and focused constituents in German. *JPh* **36**. 680–703.
- Féry, Caroline & Hubert Truckenbrodt (2005). Sisterhood and tonal scaling. *Studia Linguistica* **59**. 223–243.
- Fougeron, Cécile & Patricia A. Keating (1997). Articulatory strengthening at edges of prosodic domains. *JASA* **101**. 3728–3740.
- Goldsmith, John A. (ed.) (1995). *The handbook of phonological theory*. Cambridge, Mass. & Oxford: Blackwell.
- Grabe, Esther (1998). *Comparative intonational phonology: English and German*. PhD dissertation, University of Nijmegen.
- Grice, Martine & Stefan Baumann (2002). Deutsche Intonation und GToBI. *Linguistische Berichte* **191**. 267–298.
- Grice, Martine, Stefan Baumann & Ralf Benz Müller (2005). German intonation in autosegmental-metrical phonology. In Jun (2005). 55–83.
- Grice, Martine, Stefan Baumann & Nils Jagdfeld (2009). Tonal association and derived nuclear accents: the case of downstepping contours in German. *Lingua* **119**. 881–905.
- Gussenhoven, Carlos (1983). Focus, mode and the nucleus. *JL* **19**. 377–417.
- Gussenhoven, Carlos (1992). Sentence accents and argument structure. In Iggy Roca (ed.) *Thematic structure: its role in grammar*. Berlin & New York: Foris. 79–106.
- Gussenhoven, Carlos (2000). The lexical tone contrast of Roermond Dutch in Optimality Theory. In Merle Horne (ed.) *Prosody: theory and experiment*. Dordrecht: Kluwer. 129–167.
- Gussenhoven, Carlos (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Hart, Johan 't, René Collier & Antonie Cohen (1990). *A perceptual study of intonation: an experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press.
- Inkelas, Sharon (1989). *Prosodic constituency in the lexicon*. PhD dissertation, Stanford University.
- Jun, Sun-Ah (ed.) (2005). *Prosodic typology: the phonology of intonation and phrasing*. Oxford: Oxford University Press.
- Jun, Sun-Ah (ed.) (2014). *Prosodic typology II: the phonology of intonation and phrasing*. Oxford: Oxford University Press.
- Kabagema-Bilan, Elena, Beatriz López-Jiménez & Hubert Truckenbrodt (2011). Multiple focus in Mandarin Chinese. *Lingua* **121**. 1890–1905.
- Katz, Jonah & Elisabeth Selkirk (2011). Contrastive focus vs. discourse-new: evidence from phonetic prominence in English. *Lg* **87**. 771–816.
- Kent, R. D. & R. Netsell (1971). Effects of stress contrasts on certain articulatory parameters. *Phonetica* **24**. 23–44.
- Kentner, Gerrit & Caroline Féry (2013). A new approach to prosodic grouping. *The Linguistic Review* **30**. 277–311.
- Ladd, D. Robert (1983b). Even, focus, and normal stress. *Journal of Semantics* **2**. 157–170.
- Ladd, D. Robert (1986). Intonational phrasing: the case for recursive prosodic structure. *Phonology Yearbook* **3**. 311–340.
- Ladd, D. Robert (1988). Declination 'reset' and the hierarchical organization of utterances. *JASA* **84**. 530–544.

- Ladd, D. Robert (1990). Metrical representation of pitch register. In John Kingston & Mary E. Beckman (eds.) *Papers in laboratory phonology I: between the grammar and physics of speech*. Cambridge: Cambridge University Press. 35–57.
- Ladd, D. Robert (1993). On the theoretical status of ‘the baseline’ in modelling intonation. *Language and Speech* 36. 435–451.
- Ladd, D. Robert (2008). *Intonational phonology*. 2nd edn. Cambridge: Cambridge University Press.
- Ladd, D. Robert & Astrid Schepman (2003). ‘Sagging transitions’ between high pitch accents in English: experimental evidence. *JPh* 31. 81–112.
- Laniran, Yetunde O. & G. N. Clements (2003). Downstep and high raising: interacting factors in Yoruba tone production. *JPh* 31. 203–250.
- Lieberman, Mark & Janet B. Pierrehumbert (1984). Intonational invariance under changes in pitch range and length. In Mark Aronoff & Richard T. Oehrle (eds.) *Language sound structure*. Cambridge, Mass.: MIT Press. 157–233.
- Nespor, Marina & Irene Vogel (1986). *Prosodic phonology*. Dordrecht: Foris.
- Odden, David (1995). Tone: African languages. In Goldsmith (1995). 444–475.
- Petrone, Caterina, Hubert Truckenbrodt, Caroline Wellmann, Julia Holzgreffe, Isabell Wartenburger & Barbara Höhle (2014). On the production and perception of prosodic cues in bracketed lists. Ms, Centre for General Linguistics (ZAS), Berlin, Laboratoire Parole et Langage, CNRS UMR 7309 & University of Potsdam.
- Pierrehumbert, Janet B. (1980). *The phonology and phonetics of English intonation*. PhD dissertation, MIT.
- Pierrehumbert, Janet B. & Mary E. Beckman (1988). *Japanese tone structure*. Cambridge, Mass.: MIT Press.
- Pierrehumbert, Janet B. & Julia Hirschberg (1990). The meaning of intonational contours in the interpretation of discourse. In Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds.) *Intentions in communication*. Cambridge, Mass.: MIT Press. 271–311.
- Potts, Christopher (2005). *The logic of conventional implicatures*. Oxford & New York: Oxford University Press.
- Prieto, Pilar, Chilin Shih & Holly Nibert (1996). Pitch downtrend in Spanish. *JPh* 24. 445–473.
- Selkirk, Elisabeth (1980). Prosodic domains in phonology: Sanskrit revisited. In Mark Aronoff & Mary-Louise Kean (eds.) *Juncture*. Saratoga: Anma Libri. 107–129.
- Selkirk, Elisabeth (1986). On derived domains in sentence phonology. *Phonology Yearbook* 3. 371–405.
- Selkirk, Elisabeth (1995a). Sentence prosody: intonation, stress, and phrasing. In Goldsmith (1995). 550–569.
- Selkirk, Elisabeth (1995b). The prosodic structure of function words. In Jill Beckman, Laura Walsh Dickey & Suzanne Urbanczyk (eds.) *Papers in Optimality Theory*. Amherst: GLSA. 439–469.
- Selkirk, Elisabeth (2005). Comments on intonational phrasing in English. In Sónia Frota, Marina Vigário & Maria João Freitas (eds.) *Prosodies: with special reference to Iberian languages*. Berlin & New York: Mouton de Gruyter. 11–58.
- Selkirk, Elisabeth (2009). On clause and intonational phrase in Japanese: the syntactic grounding of prosodic constituent structure. *Gengo Kenkyu* 136. 35–73.
- Selkirk, Elisabeth (2011). The syntax–phonology interface. In John A. Goldsmith, Jason Riggle & Alan C. L. Yu (eds.) *The handbook of phonological theory*. 2nd edn. Oxford: Wiley-Blackwell. 435–484.
- Shih, Chilin (2000). A declination model of Mandarin Chinese. In Antonis Botinis (ed.) *Intonation: analysis, modelling and technology*. Dordrecht: Kluwer. 243–268.
- Truckenbrodt, Hubert (1995). *Phonological phrases: their relation to syntax, focus, and prominence*. PhD dissertation, MIT.

- Truckenbrodt, Hubert (1999). On the relation between syntactic phrases and phonological phrases. *LI* **30**. 219–255.
- Truckenbrodt, Hubert (2002). Upstep and embedded register levels. *Phonology* **19**. 77–120.
- Truckenbrodt, Hubert (2004). Final lowering in non-final position. *JPh* **32**. 313–348.
- Truckenbrodt, Hubert (2005). A short report on intonation phrase boundaries in German. *Linguistische Berichte* **203**. 273–296.
- Truckenbrodt, Hubert (2007a). The syntax–phonology interface. In Paul de Lacy (ed.) *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press. 435–456.
- Truckenbrodt, Hubert (2007b). Upstep on edge tones and on nuclear accents. In Carlos Gussenhoven & Tomas Riad (eds.) *Tones and tunes*. Vol. 2: *Experimental studies in word and sentence prosody*. Berlin & New York: Mouton de Gruyter. 349–386.
- Truckenbrodt, Hubert (2012). Semantics of intonation. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.) *Semantics: an international handbook of natural language meaning*. Vol. 3. Berlin: de Gruyter Mouton. 2039–2069.
- Truckenbrodt, Hubert (2015). Intonation phrases and speech acts. In Marlies Kluck, Dennis Ott & Mark de Vries (eds.) *Parenthesis and ellipsis: cross-linguistic and theoretical perspectives*. Berlin: De Gruyter Mouton. 301–349.
- Uhmann, Susanne (1991). *Fokusphonologie: eine Analyse deutscher Intonationskonturen in Rahmen der nicht-linearen Phonologie*. Tübingen: Niemeyer.