Upstep on edge tones and on nuclear accents

Hubert Truckenbrodt

Abstract

Southern varieties of German show a cross-linguistically unexpected upstep phenomenon. Following downstep on prenuclear peaks in an intonation phrase, upstep occurs on the nuclear peak and/or on the H edge tone of the intonation phrase. The tonal height of the upstepped peaks is at approximately the height of the first peak of the intonation phrase. Upstep occurs in non-final intonation phrases that are followed by partial reset in the following intonation phrase. Upstepped tones are argued to be scaled to a phrasal reference-line that is lowered between intonation phrases (van den Berg, Gussenhoven, and Rietveld 1992, developing a suggestion of Ladd 1988). The phrasal reference-line was first proposed to account for partial reset (a return to this reference-line in the following intonation phrase, after the reference-line is lowered). Upstep is argued to provide independent support for this reference-line in that it targets that reference-line before the reference-line is lowered. Upstep on nuclear accents is analyzed in more statistical detail in Truckenbrodt (2002). The present paper extends the empirical coverage from pitch accents to edge tones, and develops a generalized, more comprehensive account of the scaling of both pitch accents and edge tones. The results support a development of the model of scaling in which prosodic constituents mediate the assignment of phonetic reference-lines to tones in a principled way.

1. Introduction: The effect of larger prosodic domains on downstep and tonal scaling

A number of different nuclear contours were observed in utterance-internal position in the speech of 8 speakers of German. The way the high targets in these nuclear contours were pronounced motivated a careful consideration of tonal scaling in German.

Detailed empirical studies of downstep and related effects on tonal scaling are reported in Liberman and Pierrehumbert (1984) (following up on Pierrehumbert 1980) and in Ladd (1988) for English, Pierrehumbert and Beckman

(1988) for Tokyo Japanese, van den Berg, Gussenhoven, and Rietveld (1992) for Dutch, Prieto, Shih, and Nibert (1996) for Mexican Spanish, Grabe (1998) for English and German, and Connell and Ladd (1990) as well as Laniran and Clements (2003) for the tone language Yoruba.

Ladd (1988) argues that, just as adjacent accents can be downstepped, so larger adjacent domains can (and often will) also be downstepped relative to each other. Ladd implemented his suggestion in terms of hierarchically organized register features. Van den Berg, Gussenhoven, and Rietveld (1992) support this conception with data from Dutch and implement this idea using a horizontal reference-line for each such larger prosodic domain, which is lowered between the domains. Fig. 1 shows a pitch-track of a Dutch recording from van den Berg, Gussenhoven, and Rietveld (1992) onto which the author of the present paper has graphically superimposed crucial elements of their model. The plot shows downstep among accents (highlighted by superimposed black dots). At the beginning of the second prosodic domain, downstep is interrupted and a new height is targeted (*reset*). However, the new peak is not as high as the utterance-initial peak, hence this is a *partial reset*. The partial reset, in the model of Ladd (1988) and van den Berg, Gussenhoven and Rietveld (1992), is a reflection of downstep among domains. In van den Berg, Gussenhoven and Rietveld (1992), it is a return to the *phrasal reference-line* (superimposed in Fig. 1). The phrasal reference-line is lowered between domains as shown. The lowering-relation between the utterance-initial peak and the partial reset, 'downstep at a distance', is here analyzed as the result of a local downstep relation among the two larger domains.

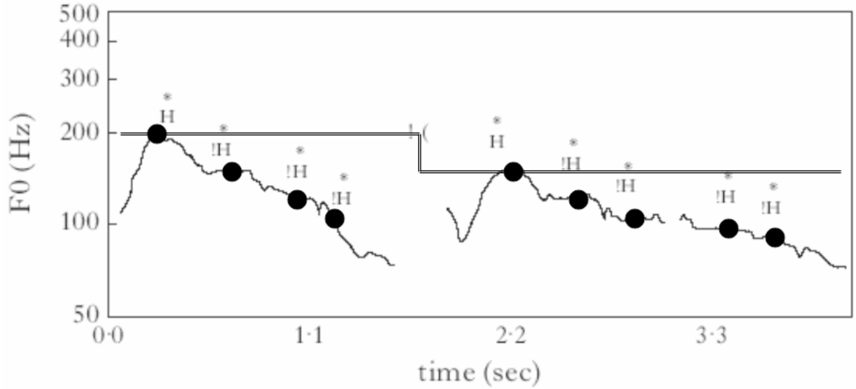
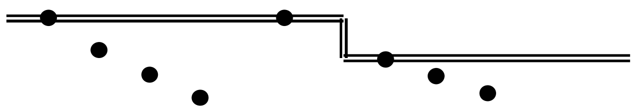


Figure 1. The Dutch utterance *(Merel, Nora, Leo, Remy), en (Nelie, Mary, Leendert, Mona en Lorna).* From van den Berg, Gussenhoven, and Rietveld (1992), with dots (for phonetic values of the H* tones) and a reference-line of the model of these authors superimposed.

Truckenbrodt (2002) reports that some speakers of German show an unexpected phenomenon of upstep on the nuclear pitch accent of an intonation phrase that is followed by another intonation phrase in the same utterance. It is argued there that these upstepped H tones are scaled to the phrasal reference-line before the phrasal reference-line is lowered, as shown in (1).

(1) [upstep] [partial reset]



That suggestion captures the core properties of upstep that are established in detail in Truckenbrodt (2002). First, the upstep shares with the (partial) reset that its height is defined independently of preceding downstep.[1] Second, the size of the upstep is not arbitrary, but is determined by the value of the initial peak. Third, the upstep is systematically higher than the following partial reset. This is captured in (1), since upstep returns to the phrasal reference-line before the end of the first intonation phrase, and thus before the reference-line is lowered. Partial reset, on the other hand, returns to the phrasal reference-line initially in the second intonation phrase, and thus after the reference-line is lowered.[2]

Independent motivation for Ladd's analysis of partial reset as the effect of downstep among larger domains is provided by the investigation of multiply embedded domains with English speakers in Ladd (1988). The main results are reproduced with German speakers in Truckenbrodt and Féry (2005) and Féry and Truckenbrodt (2005). The experiment reported there also investigates upstep in these more complex configurations. The results further confirm Ladd's analysis of partial reset as well as its implementation in terms of reference-lines by van den Berg, Gussenhoven, and Rietveld (1992) and the analysis of upstep as targeting the phrasal reference-line of that model.

The present paper draws on a larger set of recordings that were evaluated for specific aspects in Truckenbrodt (2002, 2004). The present paper presents further empirical findings from these recordings about the scaling of pitch accents and adds findings in regard to the scaling of edge tones. These motivate a more comprehensive account of the scaling of both pitch accents and edge tones. The question at the core of the account is what factors decide whether a tone is scaled on a downstepped reference-line or on the higher reference-line (upstepped). A central claim of the account is that prosodic constituents mediate between the phonological association of tones and the phonetic reference-lines (downstepped or upstepped) on which tones receive their phonetic values.

The analysis of the data and the account employ an analysis of German prosodic structure which involves two levels of phrasing, the *accentual phrase* (*A*) and the *intonation phrase* (*I*), as shown in Fig. 2. At the level of the accentual phrase, each accentual phrase carries a pitch accent. The accentual phrases and positions of accents can be predicted from focus and syntactic structure by the *Sentence Accent Assignment Rule* (*SAAR*) of Gussenhoven (1983, 1992).[3] In Fig. 2, the pitch accents are L*+H in non-final position and H+L* in final position. Downstep can be observed between the H tones of non-final L*+H pitch accents, while the sequence of the final two pitch accents, L*+H H+L* defines a more or less horizontal plateau that does not involve downstep between one H tone and the next. In the example at hand, the utterance is a single intonation phrase, as shown. The strongest stress of the intonation phrase is on the final accent assigned by the SAAR. It here falls on the preverbal object that has not undergone scrambling, which is compatible with a body of literature about the position main stress in German clauses.[4]

The analysis of the intonation contour follows the framework of Pierrehumbert (1980) and Beckman and Pierrehumbert (1986), in which the sentence melody is analyzed in terms of a sequence of H and L tones, each corresponding to a phonetic point. The F_0-contour results from interpolation among these points, so that a turning point in the F_0-contour (exempting microprosodic influences) provides evidence for the presence of a tone. The framework postulates pitch accents that associate with stressed syllables, such as L*+H in Fig. 2. The star diacritic marks the tone most closely associated with the stressed syllable and the plus sign adds a further tone to bitonal pitch accents. The framework further postulates edge tones that mark an edge of a prosodic constituent, like the final L tone in Fig. 2. The diacritic marking of edge tones here employs subscripting of the relevant domain, following Hayes and Lahiri (1991); thus L_I is an edge tone of the intonation phrase I. See Ladd (1996) and Gussenhoven (2004) for applications of this framework to a wide variety of languages.

Experimentally based applications of this framework to German intonation were first developed in Uhmann (1991) and Féry (1993). Detailed investigations of the course of downstep are reported in Grabe (1998); Grabe also offers a comparison of German and English intonation. Suggestions for a unified transcription standard for intonation contours in German (GToBI) can be found in Grice and Baumann (2002) and Grice, Baumann, and Benzmüller (2005). Studies of the intonation patterns of dialects include Fitzpatrick-Cole (1999, Bern Swiss German), Barker (2002, Tyrolean German), Peters (2002, Hamburgisch), Atterer and Ladd (2004, comparing prenuclear rises of Northern and Southern German speakers), Kügler (2005, Swabian and Upper Saxon German) and Gilles (2005).

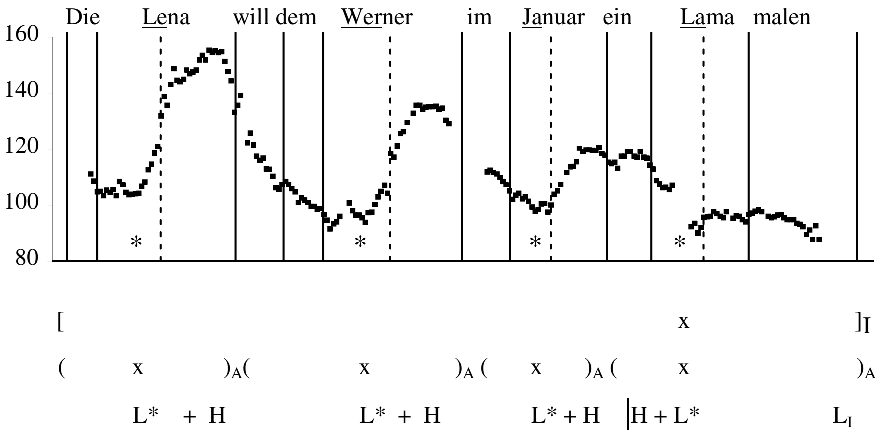


Figure 2. 'Lena wants to paint a llama for Werner in January.' Solid lines delimit words; dashed lines delimit accented syllables, which are marked by a star. Speaker TL. Adapted from Truckenbrodt (2004).

The material discussed in this paper extends the tonal analysis of Truckenbrodt (2002, 2004). Speakers from Southern regions were recorded (Baden-Württemberg and Austria). The predominant tonal pattern found in single clauses is the one illustrated in Fig. 2, with L*+H on non-final pitch accents and H+L* on final pitch accents. The stressed syllable is perceived as low in both instances. The L*+H pitch accent is also reported for Southern dialects in Fitzpatrick-Cole (1999), Barker (2002), and Kügler (2005). Both L*+H and H+L* are in the inventory of German pitch accents suggested in Grice and Baumann (2002) and Grice, Baumann, and Benzmüller (2005). The final H+L* of Fig. 2 is documented in more detail in Truckenbrodt (2004).

2. Summary of methods

A summary of the methods is given here. For additional details, the reader is referred to the methods sections in Truckenbrodt (2002, 2004). The stimuli include a one-clause condition from which Fig. 2 is drawn. The crucial element is the two-clause condition, in which upstep can be observed on different tones at the end of the first clause. The data for the two-clause condition were recorded in four blocks, whose prosodic make-up is schematically shown in (2). The two clauses are coordinated with *and*. Following a cross-linguistic generalization (Nespor and Vogel 1986), each of these coordinated clauses was expected (and found) to form a separate intonation phrase. These are indicated by square brackets in (2). In the first clause/intonation phrase, the four series range in prosodic length from two accents in the shortest series (a) to five accents in the longest series (d). In the second clause, the constant length is three accents.[5] Single underlining in (2) highlights accented elements, double underlining highlights the nuclear (last and assumed strongest) stress of each intonation phrase.

(2)

		accents in first clause
a.	[SU Aux DO V] [XP XP XP]	2
b.	[SU & SU Aux DO V] [XP XP XP]	3
c.	[SU & SU Aux IO DO V] [XP XP XP]	4
d.	[SU & SU Aux IO PP DO V] [XP XP XP]	5

(SU: subject, SU & SU: subject coordination, Aux: finite auxiliary verb, IO: indirect object, PP: prepositional phrase adjunct, DO: direct object, V: infinitive main verb, XP: syntactic phrase)

An effort was made to use sonorant sounds in the stimuli as much as possible, in particular in the positions of expected H tones. A list of all stimuli with their transcriptions can be found in Truckenbrodt (2004). One example stimulus from each set a–d is shown in (3), here with the expected prosodic structure indicated; this indication was not part of the presentation of the stimuli to the subjects.

(3) a. [*(der Maurer)(will das Weben lernen)] [(und die Hanne)(soll ihm Leinen)(und Wolle besorgen)*]
'The bricklayer wants to learn weaving, and Hanne is supposed to get linen and wool for him.'

b. [*(der Werner)(und die Lola)(wollen malen lernen)] [(und die Manu)(will dem Lehrer)(Romme zeigen)*]
'Werner and Lola want to learn painting, and Manu wants to show rummy to the teacher.'

c. [*(der Maler)(und der Lehrer)(wollen der Hanne)(Maronen geben)] [(und der Maurer) (will der Lena)(Murnau zeigen)*]
'The painter and the teacher want to give chestnuts to Hanne, and the bricklayer wants to show Murnau to Lena.'

d. [*(der Maurer)(und sein Lehrling)(wollen dem Werner)(in Kamerun)(ein Lama malen)] [(und der Maler)(will im Januar)(in Murnau wohnen)*]
‘The bricklayer and his apprentice want to paint a llama for Werner in Cameroon, and the painter wants to live in Murnau in January.’

Of the eight speakers that are analyzed, six were from Baden-Württemberg and two from Austria (see Truckenbrodt 2002 for more details).

The speakers read the sentences from a pseudo-randomized list. To control for focus, each sentence was read as an answer to the question *‘was gibt’s Neues?’*, ‘what’s new?’ Brief instructions clarified that renditions in a normal voice were required, without any emphatic highlighting of any element or the expression of particular emotions, like surprise.

Each set a–d in (2) was represented by six stimulus sentences, and each of these was read three times by each speaker. Therefore, each set in (2) is represented by 18 token recordings from each speaker. The recordings were analyzed with ESPS/waves+. A tonal analysis was hypothesized for each token. The speakers assigned L*+H very consistently in the expected positions of non-final pitch accents, and each speaker was very consistent in his/her assignment of boundary tones preceding the medial intonation phrase break. A prevalent tonal pattern could thus be postulated for each speaker, with very few exceptions that were mostly limited to the utterance-final pitch accent. The exceptions were not included in the measurements.

The turning points that are evidence for the tones of the analysis were measured according to fixed criteria. In a few cases, measurements had to be skipped, which was also done according to fixed criteria, for which see Truckenbrodt (2002, 2004). The average values of the measurements that are presented below for the different speakers are thus based on the prevalent tonal pattern used by a speaker, and present, for each tone of this prevalent tonal pattern, averages of its phonetic height in the recordings.

The stimuli of the one-clause condition and of the two-clause condition were recorded during one recording session. In a second recording session, a few days after the first, further details of the intonation were tested with different stimuli. From this second session, the present paper includes the results of six readings by each speaker of the sentence in (4), here shown with its prosodic structure.

(4) [*(der Maurermangel)(und der Lehrermangel)(werden der Regierung) (Probleme bereiten)*]
'The lack of bricklayers and the lack of teachers will pose problems for the government.'

In this sentence, the compound stress in the first and second accentual phrase leads to accents in preantepenultimate position of their respective accentual phrases. This allows for the effect of a boundary tone at the right edge of the first two accentual phrases to emerge.

3. Earlier reports of results from this data concerning pitch accents

Four of the eight speakers showed upstep on the L*+H nuclear pitch accent of the first intonation phrase. The properties of upstep of these four speakers are analyzed in detail in Truckenbrodt (2002). The main results are as reviewed in connection with (1) above. These speakers are integrated into the discussion below (they constitute what is here called Pattern 3 and Pattern 4). Truckenbrodt (2004) investigates the exact course of downstep among the pitch accents in this data, with particular attention to final lowering (Liberman and Pierrehumbert 1984). The basis of that evaluation are the data of the single-clause condition, and the data of the two-clause condition of the four speakers showing upstep on the nuclear pitch accent. The main result is that peaks of pitch accents followed by downstep are higher than peaks of pitch accents that are not followed by downstep. This is briefly related to the findings reported here in section 6.3.

The present paper provides a comprehensive discussion of the two-clause condition for all eight speakers. It includes the scaling of edge tones in addition to the scaling of pitch accents. A tonal analysis is developed that brings out the important common elements among the eight speakers and reduces their main difference to phonological surface variation. The tonal analysis is the basis for a general account of the scaling of edge tones and pitch accents across single- and two-clause condition.

4. Results and phonological analysis

4.1. H_A edge tones of the accentual phrase

In this section, the recordings of the sentence in (4) are evaluated. It provides the background to the understanding of the tonal analysis of the two-clause condition. A pitch-track by speaker TL is shown in Fig. 3.

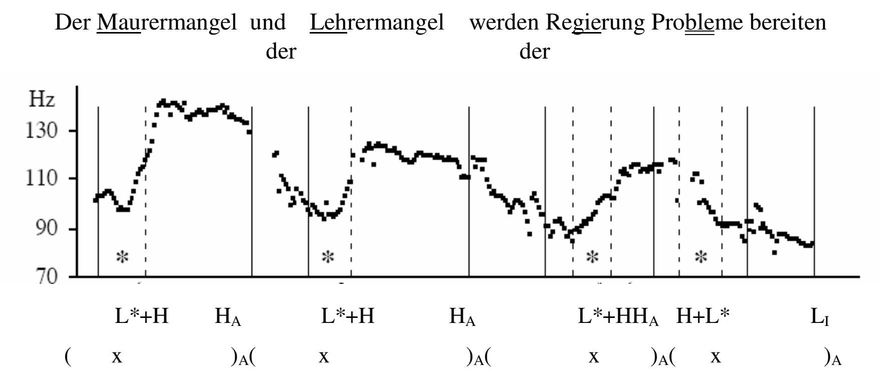


Figure 3. 'The lack of bricklayers and the lack of teachers will pose problems for the government.' Solid lines delimit accented words; dashed lines delimit accented syllables, which are marked by a star. Speaker TL.

In both the first and the second accentual phrases, a horizontal plateau (here slightly falling) can be observed spanning the distance from the end of the L*+H rise to the end of the accentual phrase. Such a plateau is observed more generally, as can be seen in the averaged measurements in Fig. 4.

The plateau is evidence for the presence of a H_A edge tone of the accentual phrase at the end of the plateau. The full tonal specification of non-final accentual phrases is thus L*+H H_A. The regular presence of H_A in this position in turn provides evidence for the prosodic constituent accentual phrase.

A further point of interest concerns the scaling of H_A in both the first and the second accentual phrase. The first L*+H rise is higher than the second L*+H rise due to downstep of the second rise. While the plateaus are not literally horizontal in many cases, it seems clear for each speaker that the first occurrence of H_A is scaled relative to the end of the first rise, and that the second occurrence of H_A is scaled relative to the end of the second, downstepped, rise. Idealizing somewhat, the account of scaling should thus assign H_A in L*+H H_A to the same height as the immediately preceding '+H' in initial position as well as in downstepped position.

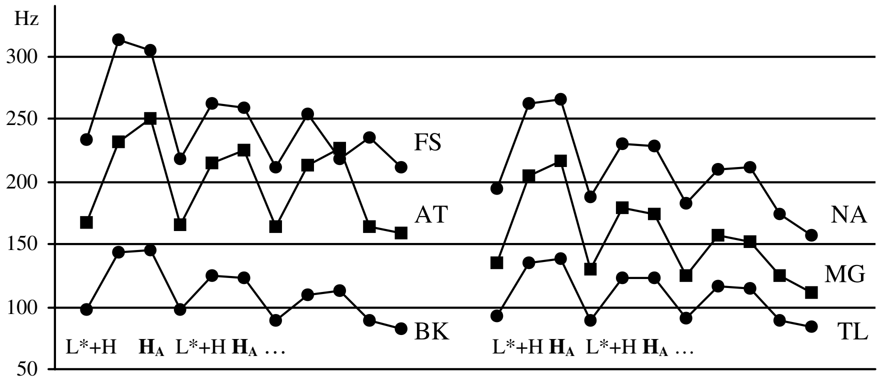


Figure 4. Averaged measurements of L and H turning points of six renditions by each speaker of the sentence *Der Maurermangel und der Lehrermangel werden der Regierung Probleme bereiten.* 'The lack of bricklayers and the lack of teachers will pose problems for the government.'

4.2. Tonal analysis in position preceding the medial intonation phrase boundary

A central interest in this paper concerns the intonational consequences of the intonation phrase boundary between the two clauses in the two-clause condition.

The speakers are here grouped into four patterns depending on how they mark the medial intonation phrase break. Patterns 1 and 4 are each represented by a single speaker from Austria. Patterns 2 and 3 are each represented by three speakers from Baden-Württemberg. Because of the small sample, no claims are made about regional dialects or representativeness of the four patterns found. However, since each speaker was very coherent in the marking of the intonation phrase boundary across the 72 recordings of the two-clause condition, the patterns are more than accidental renditions of individual utterances.

In the following evaluations of the four patterns, a phonological analysis is assumed that stems from the underlying tonal structure in (5). Thus, the L*+H H_A pattern of non-final accentual phrases (as established in the preceding section) is also found in non-final position before the medial intonation phrase boundary. To this a following H_I tone is added, marking continuation, at the end of the non-final intonation phrase. The analysis in terms of a sequence of two edge tones, with the one marking the lower edge preceding the one marking the higher edge is similar to the analysis of English in Beckman and Pierrehumbert (1986), though the prosodic domains differ in the two cases. The lower domain is here the accentual phrase (the domain of accent assignment), due to the independent evidence for the assignment of a H_A edge tone at the end of this domain, as in Fig. 3 and Fig. 4.

(5) Underlying tonal structure postulated for all patterns in position preceding the medial intonation phrase boundary:

```
          x                ]I [
... )A (  x                )A ( ...
          L* + H   HA HI
```

Of interest is the (phonological and phonetic) realization of the three H tones in (5) at the end of the non-final intonation phrase.

4.3. Nuclear pattern 1

Fig. 5 shows a typical rendition of the first clause and beginning of the second clause by speaker FS, who (alone) represents pattern 1. Here and in later illustrations, the boundary between the two clauses (and the breath pause) is highlighted by a vertical bar, and the course of F_0 in the last accentual phrase preceding this is highlighted by an oval. In the present analysis, the underlying tones of (5) surface unchanged in pattern 1. Thus, there is a nuclear L*+H rise, of which the H is downstepped relative to a preceding downstep. Following this, there is evidence for a further H tone, here analyzed as H_A, slightly lower than the end of the preceding rise. Finally, there is evidence for the H_I tone of the analysis in the high rise up to the point where the breath-pause begins.

Fig. 6 shows averaged values of the four sets of different lengths from this speaker. The plot shows the very coherent downstepping pattern in the first clause that includes the nuclear accent of the first clause. It further shows that the H_A-values of the four series (plotted at a single point H_A) are always lower than the ends of the preceding nuclear rises. (Relative to the preceding nuclear L*+H rise, the difference in height between L* and H_A is on average .64; the height difference between H_A and +H is significant by a Paired sample t-test ($t = 13.9[df= 71]$, $p <.001$). Further, the final H_I values reach a comparable height across the four sets, which is higher than the height of the utterance-initial rises. (The difference between +H of the utterance-initial rise and H_I is significant by a Paired sample t-test; $t = -9.1[df=71]$, $p < .001$). There is partial reset initially in the second clause: These values are higher than the end of the long downstep sequences in the first clause, yet lower than the utterance-initial peak.

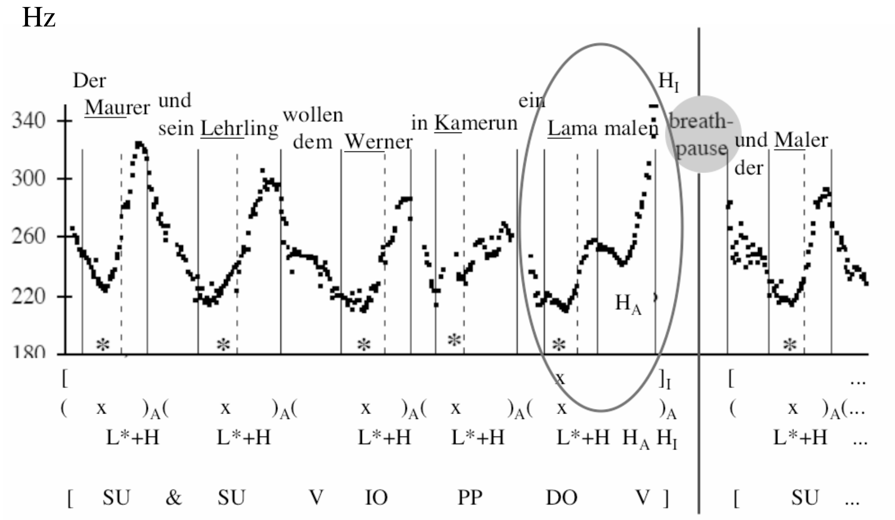


Figure 5. Illustration of pattern 1. 'The bricklayer and his apprentice want to paint a llama for Werner in Cameroon, and the painter (… = wants to live in Murnau in January).' Solid lines delimit accented words; dashed lines delimit accented syllables, which are marked by a star. Speaker FS.

The phonetic analysis accounts for the values of H_A and H_I as an overlay of two factors of scaling: the target-value of H_A will be that of the preceding downstepped H of the nuclear L*+H. The target-value of H_I will be that of the utterance-initial (L*+)H; they are thus analyzed as upstepped as in the model in (1). Overlaid is a postulated phonetic dissimilation effect among the two edge tones H_A and H_I, moving them apart from each other in their phonetic height. Therefore H_A is lower than its abstract target, and H_I is higher than its abstract target. The postulated phonetic dissimilation apart, then, we can schematically represent this pattern as in (6). The tonal value that is analyed as upstepped is highlighted with a dot.

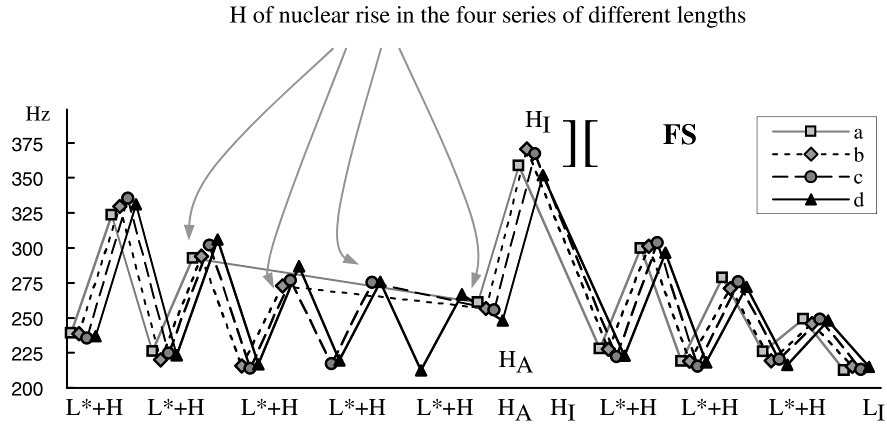
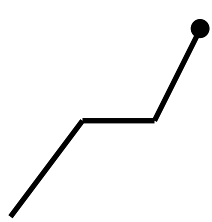


Figure 6. Averaged values of speaker FS, pattern 1. The four series begin with downstepping L*+H pitch accents (two of them in series a., three in series b., four in series c., and five in series d.). The last (nuclear) rise in each series is highlighted with an arrow reaching from the top. In each series, it is followed by the edge tones H_A (plotted at around the same x-axis point) and H_I (likewise). After the intonation phrase boundary, there follow three downstepping L*+H pitch accents and a final L_I edge tone (in all series). Each plotting point is an average of 18 measurements minus any missing values; there are no missing H values in the first clause.

(6) Nuclear pattern 1

L*+H H_A H_I (with all underlying tones preserved)



4.4. Nuclear pattern 2

Pattern 2 is represented by three speakers. It is minimally different from pattern 1: In pattern 1, the speaker shows consistent evidence of three H tones following the nuclear L*: a local maximum analyzed as '+H', a local minimum analyzed as H_A, and a very high point just at the intonation phrase break, analyzed as H_I (see Figs. 5 and 6). In pattern 2, there is consistently evidence for only two H tones following the nuclear L*, as illustrated in Fig, 7: one point at which a steep rise turns into a more shallow rise in the recordings, analyzed as '+H', and a very high point at the intonation phrase boundary, analyzed as H_I. The nuclear contour of pattern 2 is analyzed, on the surface, as L*+H H_I. The underlying H_A tone is taken to be deleted in pattern 2.[6]

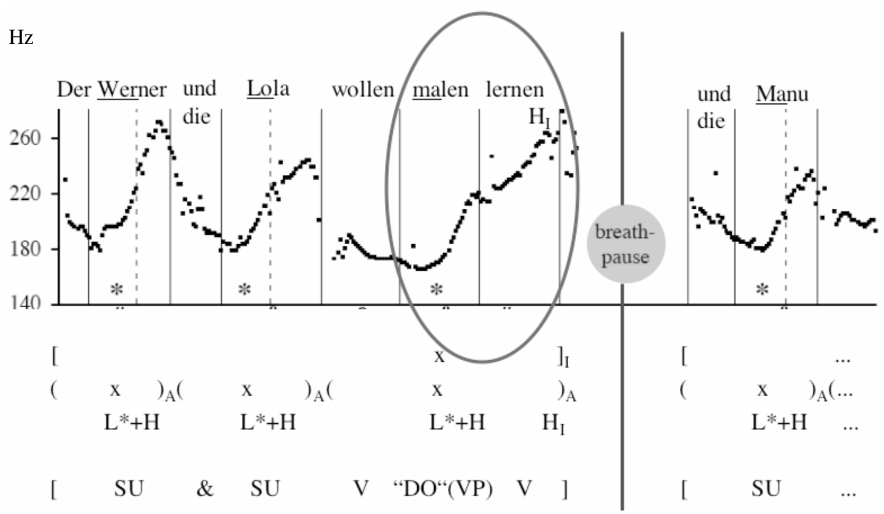


Figure 7. Illustration of pattern 2. 'Werner and Lola want to learn to paint, and Manu (…= wants to show rummy to the teacher).' Solid lines delimit accented words; dashed lines delimit accented syllables, which are marked by a star. Speaker NA.

Fig. 8 shows the plots of all four sets from the three speakers of pattern 2. The tone H_I can be seen to be upstepped to a height comparable to the utterance-initial peak, allowing for some amount of declination. With H_A deleted, no raising of H_I by phonetic dissimilation from H_A is expected. Calculated in relation to the utterance-final L_I tone[7], the height of H_I is on average .92 of the height of the utterance-initial +H for speaker AT; the ratio is .96 for speaker BK and .91 for speaker NA. The difference between the utterance-initial +H and H_I are significant by a Paired sample t-test for each of the three speakers (*speaker AT*: $t = 4.7[df=71]$, $p < .001$; *speaker BK*: $t = 3.2[df=71]$, $p < .01$; *speaker NA*: $t = 5.8[df=64]$, $p < .001$).

The slightly lower values of H_I relative to the utterance-initial rises are plausibly attributed to declination across the utterance; see Pierrehumbert 1980;

Gussenhoven and Rietveld 1988; Ladd 1988, 1993; Grabe 1998; Shih 2000. With this, the data from pattern 2 support an analysis of H_I in this pattern as

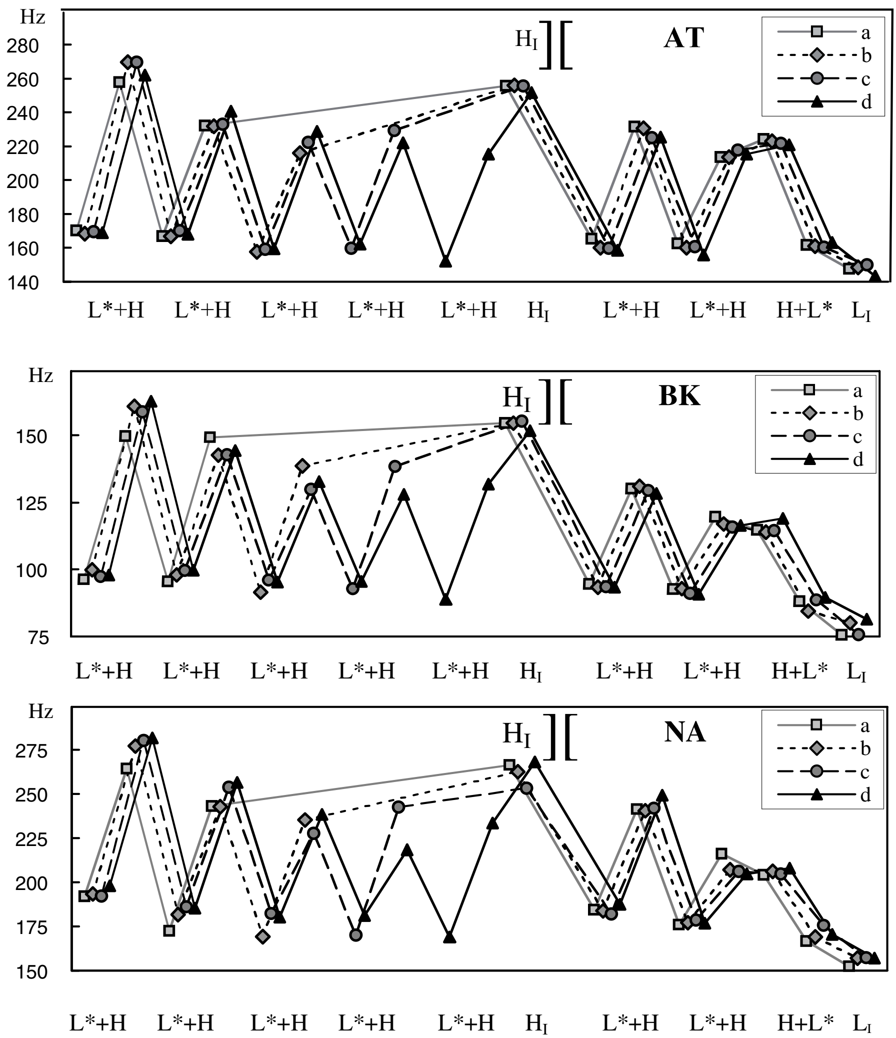


Figure 8. Averaged values of the three speakers of pattern 2. The four series begin with downstepping L*+H pitch accents. The last (nuclear) one of these in the first clause is downstepped in some, but not all of the series. In each series, it is followed by the edge tone H_I (plotted at around the same x-axis point). After the intonation phrase boundary, there follow two downstepping L*+H pitch accents and a final sequence H+L* L_I (in all series). Each plotting point is an average of 18 measurements minus any missing values; the missing H values (out of 324 H values) in the first clause are, for AT: 13/324; for BK: none; for NA: 42/324.

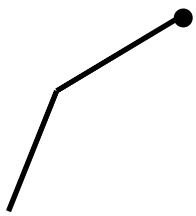
upstepped as in (1). The data in Fig. 8 are compatible with the aspect of (1) by which the initial peak in the second clause is in partial reset, even though the evidence for the reset is not strong in pattern 2. Renewed height in the initial peak of the second clause relative to downstep in the first clause is really discernible only for speaker NA.

The scaling of the nuclear rise in the first clause varies among speakers and series in ways that are not entirely clear. It is lower than the preceding peak in half of the series of the three speakers, but of equal or slightly raised height in the other series. The speakers are not consistent in their relative scaling of the nuclear rise across the four series. On the whole, the nuclear rise is clearly lower than both utterance-initial peak and H_I. For concreteness, it is here classified as downstepped relative to the preceding peak (rather than upstepped to the initial height), with the assumption that another effect of scaling (perhaps boosting due to prominence) may sometimes have a small overlaid effect of raising on it.

We can schematically represent pattern 2 as in (7). The dot again highlights the value analyzed as upstepped.

(7) Nuclear pattern 2

L*+H $\cancel{H}_A$ H_I (with /H_A/ deleted before H_I)



4.5. Nuclear pattern 3

Pattern 3 is represented by three speakers. This pattern also shows no evidence of H_A, and H_A is also taken to be deleted here. Pattern 3 brings in a new element in the present paper. While patterns 1 and 2 show upstep only on H_I, pattern 3 shows upstep both on H_I and on the preceding end of the rise, the H tone of the nuclear L*+H. (This upstep on the nuclear rise is evaluated in detail in Truckenbrodt 2002.) Fig. 9 shows a representative pitch track. The dramatic rise on the nuclear accent is very clear.

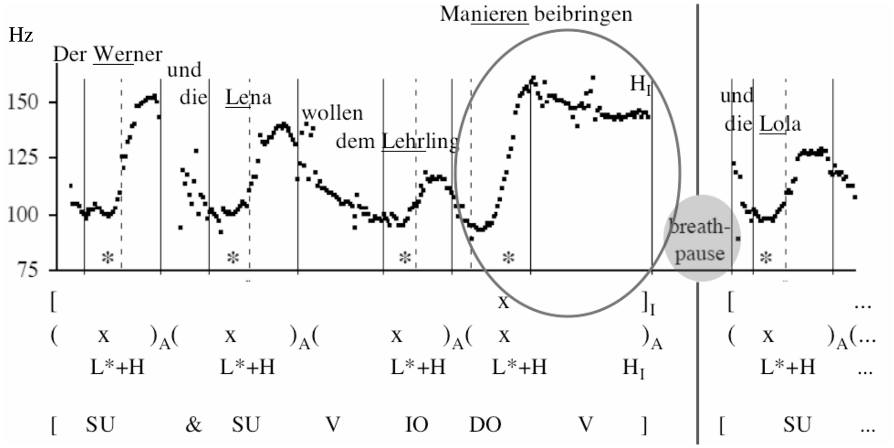


Figure 9. 'Werner and Lena want to teach manners to the apprentice, and Lola (…=wants to give a warning to Manuel).' Solid lines delimit accented words; dashed lines delimit accented syllables, which are marked by a star. Speaker TL.

Fig. 10 shows the averaged values for the same speaker of pattern 3. The ends of the nuclear rises in the four sets of different lengths are highlighted. The nuclear rise is clearly upstepped relative to preceding downstep (which is visible in particular in the longer sequences). The height of H_I, plotted at the same point of the x-axis for all four series, is comparable across series. The endpoints of the nuclear rises are of comparable height across the four sets and are each of comparable height to the following H_I values (labeled "H%" in Fig. 10) and to the height of the utterance-initial rise. With H_A deleted in the phonological analysis, no phonetic dissimilation of H_I to a height above its upstepped target (at the height of the initial rise) is expected or found.

In more detail, relative to the height of the utterance-final L_I, the nuclear upstepped +H is on average .94 of the height of the utterance-initial peak, and H_I is on average .92 of the height of the utterance-initial peak. The difference between the utterance-initial peak and each of the two upstepped H tones is significant by a Paired sample t-test ($H_{U\text{-}ini}$ *and nuclear +H*: t = 4.6[df=71], p < .001; $H_{U\text{-}ini}$ *and* H_I: t = 5.1[df=71], p < .001). This is again attributed to declination. It is compatible with this interpretation that there is no significant difference in height between the two upstepped peaks by a Paired sample t-test (*nuclear +H and* H_I: t = 1.6[df=71], p > .05). See Truckenbrodt 2002 for plots that include the upstepped nuclear rise and the upstepped H_I of the other two speakers of pattern 3, CB and MG. Speaker CB shows no effect of declination: relative to the height of the utterance-final L_I, the upstepped nuclear rise is on average .98 of the height of the utterance-initial peak, and H_I is on average

1.01 of the height of the utterance-initial peak. Pairwise comparison of these three H values by Paired sample t-tests shows that there are no significant distinctions ($H_{U\text{-}ini}$ *and nuclear +H*: t = 1.5[df=71], p > .05; $H_{U\text{-}ini}$ *and* H_I: t = 0.3[df=71], p > .05; *nuclear +H and* H_I: t = –1.8[df=71]; p > .05). Truckenbrodt (2002) argues that speaker MG shows some amount of systematic register undershoot in the upstepped values. However, for this speaker, like for TL and CB, the upstepped H_I is not significantly different in height from the upstepped nuclear +H by a Paired sample t-test (t = 1.4[df=71], p > .05).

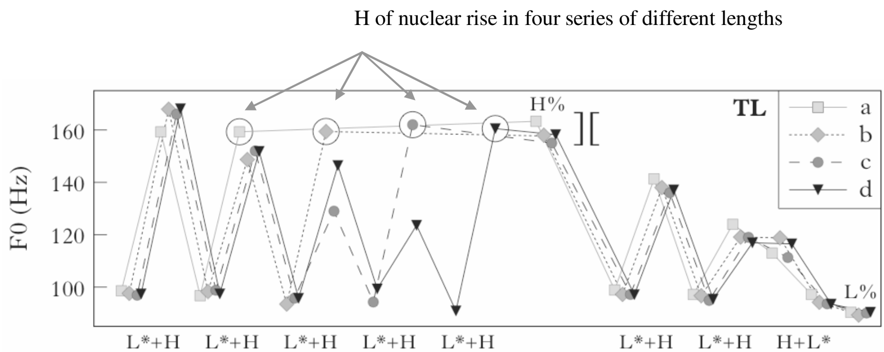


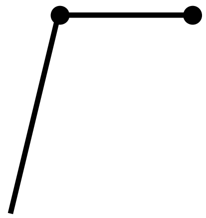
Figure 10. Averaged values of speaker TL of pattern 3. The four series begin with L*+H pitch accents. These are downstepped up to the nuclear (L*+)H tones of the first clause (circled), which are upstepped to around the initial height, undoing preceding downstep. The nuclear rise is the second rise of series a., the third rise of series b., the fourth of series c., and the fifth of series d. In each series the nuclear rise is followed by a H_I (in the plot: 'H%') tone, which is likewise upstepped. From Truckenbrodt (2002) with arrows and comment at the top added. Each point is an average of 18 measurements minus any missing values; there are 2 missing H values in the first clause, out of 324 H values in the first clause.

The data support an analysis of pattern 3 in which both the nuclear rise and the following H_I are upstepped in the model of upstep in (1), allowing for some amount of speaker-specific declination. Further support for this analysis comes from the partial reset on the initial peak of the second clause (see Fig. 10). This peak is higher than the end of the downstep sequences in the first clause, yet lower than both the utterance-initial peak and the upstepped peaks in the first clause.

Pattern 3 is schematically represented in (8), with the two upstepped H values highlighted. The upstepped scaling of the nuclear rise in this pattern is accounted for below.

(8) Nuclear pattern 3

L*+H $\cancel{H}_A$ H_I (with /H_A/ deleted before H_I)



4.6. Nuclear pattern 4

One may wonder whether pattern 3 shows some sort of leftward extension of the upstep on H_I, given that upstep was found isolated on H_I in patterns 1 and 2. There is, however, an argument for maintaining that upstep on the nuclear rise in pattern 3 is a separate phenomenon that combines with upstep on H_I in pattern 3. The argument comes from pattern 4 and is adopted here, even though this pattern is represented only by a single speaker. This speaker regularly shows upstep on the nuclear rise, as seen in Fig. 11, which cannot somehow be inherited from the right, since the following tones do not have a comparable height. The following tones here are a L and a H tone, analyzed as L_A H_I arising from underlying H_A H_I by phonological dissimilation of H_A before H_I.

The presence of these tones and their scaling are entirely regular for this speaker, as shown in Figure 12. The H tones of the nuclear rises of this speaker are again highlighted. As in pattern 3, they are of comparable height among each other, and of comparable height to the utterance-initial rises. They are clearly upstepped relative to preceding downstep, as can be seen in particular by tracing the longer sequences in Figure 12 (The upstepped +H is on average 1.04 of the height of the utterance-initial peak, with the height calculated against the utterance-final L_I; the difference between the utterance-initial +H and the upstepped +H is not significant by a Paired sample t-test ($t = -1.0$ [df=71], $p > .05$).

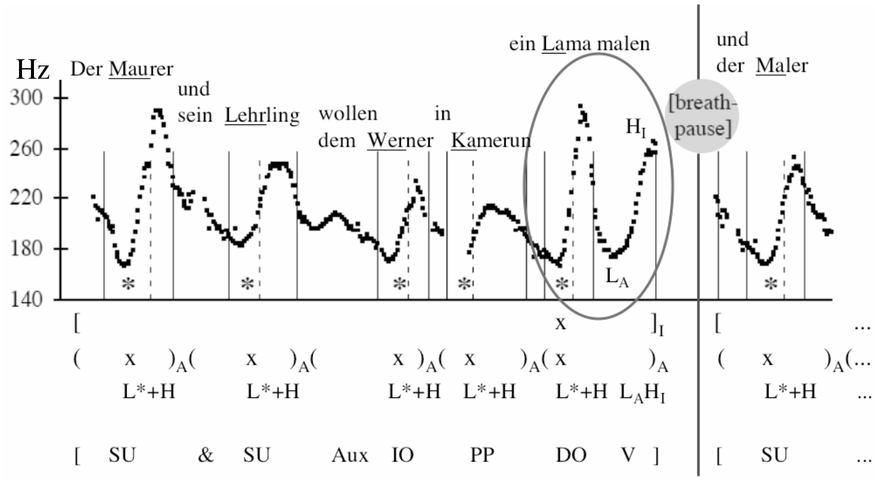


Figure 11. 'The bricklayer and his apprentice want to paint a llama for Werner in Cameroon, and the painter (…= wants to live in Murnau in January).' Solid lines delimit accented words; dashed lines delimit accented syllables, which are marked by a star. Speaker SW.

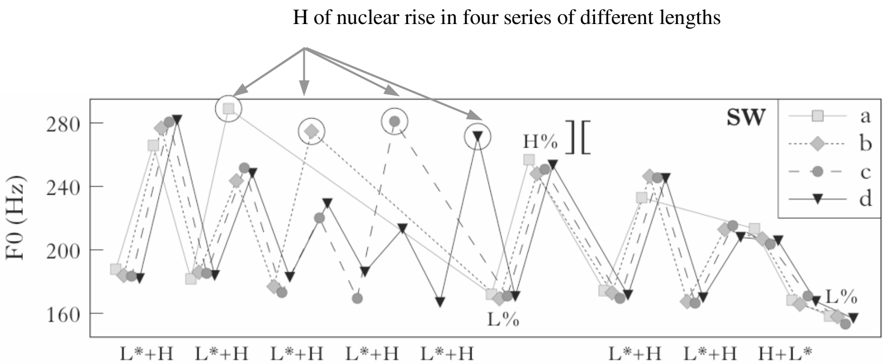


Figure 12. Averaged values of speaker SW, pattern 4. The four series begin with L*+H pitch accents. These are downstepped up to the nuclear (L*+)H tones of the first clause (circled), which are upstepped to around the initial height, undoing preceding downstep. The nuclear rise is the second rise of series a., the third rise of series b., the fourth of series c., and the fifth of series d. In each series the nuclear rise is followed by a L_A H_I sequence of edge tones (in the plot: "L%" and "H%"). From Truckenbrodt (2002) with arrows and comment at the top added. Each point is an average of 18 measurements minus any missing values; there are 2 missing H values in the first clause, out of 324 H values in the first clause.

The data thus support an analysis of the nuclear rise of pattern 4 as upstepped in the model in (1). The relation of the upstep relative to the partial reset initially in the second clause supports this interpretation. This can be assessed in Fig. 12 as well as in Fig. 13 which plots only the H values of the same data. The values in partial reset are higher than the end of the longer downstep sequences in the first clause, but lower than the utterance-initial peaks. Importantly, the upstepped nuclear peaks are also clearly higher than the values in partial reset.

The scaling the nuclear rise and of H_I for this speaker are discussed below. Pattern 4 is schematically represented in (9), with the upstepped nuclear peak highlighted.

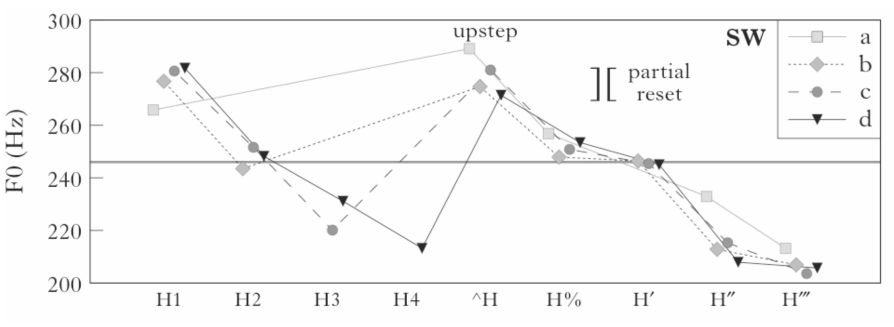
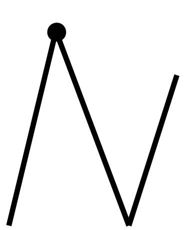


Figure 13. H values (omitting L values) of SW, pattern 4. The upstepped nuclear values are here plotted at around the same point of the x-axis, labeled ^H, and the H_I values are represented as H%. From Truckenbrodt (2002).

(9) Nuclear pattern 4

L*+H L_A H_I (with H_A dissimilated to L_A before H_I)



The upstep on the nuclear pitch accent of patterns 3 and 4 is analyzed in detail in Truckenbrodt (2002), as discussed in the introduction of the present paper.

4.7. Summary of the results on upstep and the phonological analysis

The results of the eight speakers, including nuclear pitch accents and edge tones, show that upstep on the nuclear rise is not obligatory in medial position (patterns 1 and, less clearly 2, show downstep on the nuclear rise), and that additional upstepped values are found with the H_I edge tones in patterns 1–3. These upstepped values also support the model of scaling of upstepped tones in (1). Two effects of tone scaling that seem to be overlaid on upstep are speaker-specific declination and what may be an effect of phonetic dissimilation among adjacent edge tones.

The phonological analysis derives the different tonal patterns found from the underlying sequence L*+H H_A H_I. The sequence surfaces unchanged in pattern 1. Patterns 2 and 3 are analyzed in terms of deletion of H_A before H_I, and pattern 4 in terms of dissimilation of H_A to L_A before H_I.

5. Phonetic reference-lines and prosodic levels

When is a tone scaled with the help of the downstep function applying within the phrase, and when may or must it be scaled as a return to the phrasal reference-line? In this section, a principled answer to this question is developed, in an extension of the model of Ladd (1988) and van den Berg, Gussenhoven, and Rietveld (1992). The account is also a development of related suggestions in Truckenbrodt (2002), Truckenbrodt and Féry (2005), and Féry and Truckenbrodt (2005). A core claim of the account is that prosodic constituents mediate the assignment of phonetic reference-lines to tones. This view has initial plausibility, since the phrasal reference-line is lowered once per intonation phrase, and as such can be said to be a property of the intonational phrase. The upstepped tones that are scaled to this reference-line are tones that in some sense also pertain to the level of the intonation phrase: the edge tone H_I of the intonation phrase, and the nuclear L*+H on the strongest stress of the intonation phrase. These connections enter into the more general account developed in this section.

5.1. Definitions

Reference-lines are first defined as in (10).

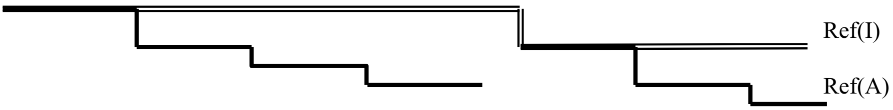
(10) Prosodic constituents and reference-lines
  a. To each prosodic level X, assign a reference-line Ref(X) for H tones and lower it once per constituent of level X.

b. For a prosodic constituent X which is initial in a higher prosodic constituent Y, Ref(X) = Ref(Y).

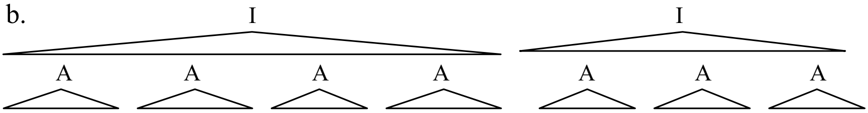
(10) takes a prosodic representation like (11b) as input, and defines reference-lines for H tones on the basis of it, in this case (11a). The reference-line of the intonation phrase I (the phrasal reference-line of van den Berg, Gussenhoven, and Rietveld 1992) is lowered once per intonation phrase I. The reference-line of the accentual phrase A is lowered once per accentual phrase A. (10b) identifies the height of the A-reference-line with the height of the I-reference-line initially in each of the two Is.

(11) Reference-lines from prosodic constituents

a.



b.



Tones are associated with prosodic levels as in (12). Edge tones H_A are associated with A, edge tones H_I are associated with I. A pitch accent can be associated with an accentual phrase because it marks the strongest stress of the accentual phrase. The nuclear pitch accent of an I may alternatively be associated with I, of which it also marks the strongest stress. (12) is inspired by, and very close to, associations between tones and prosodic constituents by Pierrehumbert and Beckman (1988), though these authors do not use these associations to condition the choice of downstepped vs. upstepped scaling as done here. (13) then provides a crucial link between tonal association as in (12) and choice of a reference-line for that tone, with reference-lines as defined in (10).

(12) **Tones and prosodic constituents**

a. Edge tones are associated with the constituent of which they mark the edge.

b. Pitch accents are associated with a constituent of which they mark the highest prominence.

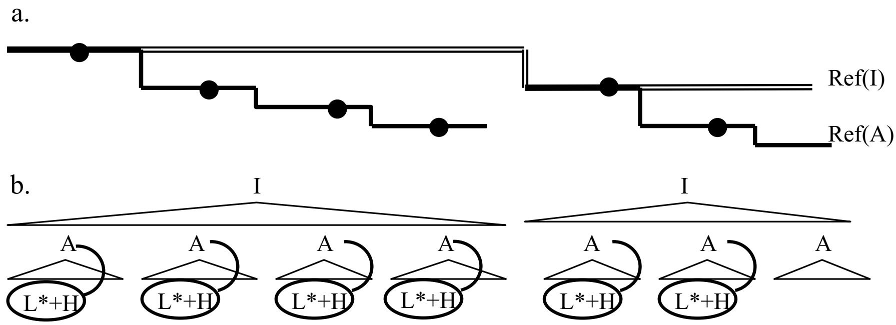
(13) **Tone scaling**

If a H tone is associated with prosodic level X, it is scaled to Ref(X).

5.2. Application to tones associated with the accentual phrase

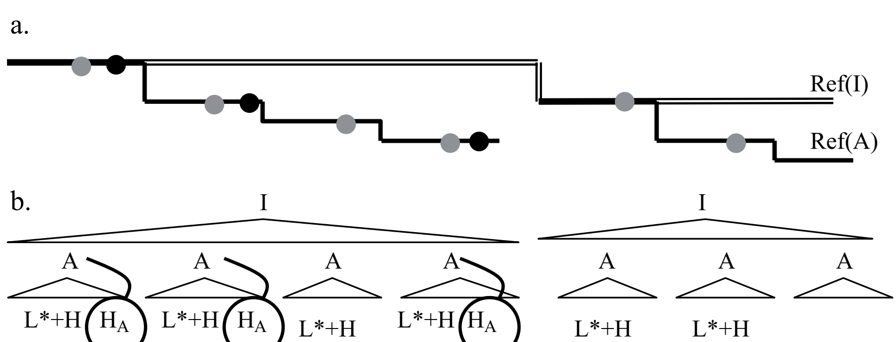
(14b) shows the pitch accents associated with accentual phrases. Each pitch accent (including the nuclear one, to which the discussion will return) marks the strongest stress of an accentual phrase A, and is associated with that A in (14b). (14a) shows the downstepped scaling of the H tones of the pitch accents predicted by (13). Since the pitch accents are associated with A, their H tones are scaled on the reference-line(s) of A.

(14) Scaling of the pitch accents associated to A



(15) shows (in black) the association and scaling of some H_A edge tones with A. (The scaling of the pitch accents is here backgrounded in grey.) As shown, (13) puts H_A in L*+H H_A at the same height as the preceding '+H' in both initial and downstepped positions. This is the desired result for the observations that were made in connection with Figs. 3 and 4.

(15) Scaling of H_A edge tones

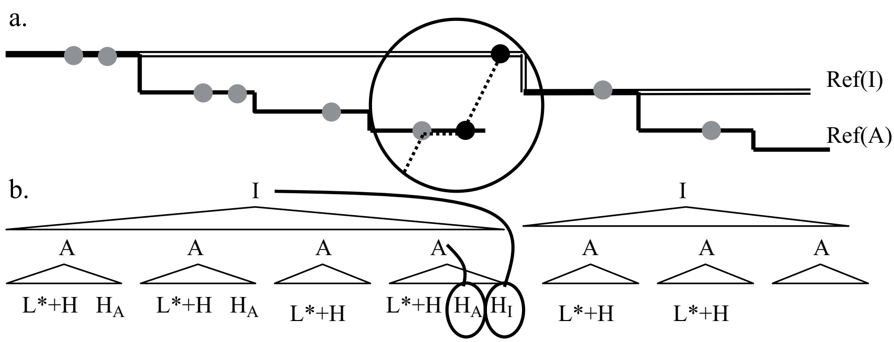


It can be seen that the present model also scales an H_A tone after a downstepped nuclear rise to the height of the preceding nuclear +H. This case occurs in pattern 1, the only pattern that does not delete or change H_A before H_I. The scaling of this H_A on the downstepped reference-line of the accentual phrase is the desired result in this case as well; compare the discussion of Figs. 5 and 6 above.

5.3. Tones associated with the intonation phrase and the account of patterns 1–3

The account of pattern 1 is then completed by consideration of edge tone H_I, associated with I and thus scaled on the reference-line of I to an upstepped height, as shown in (16); see the discussion of Figs. 5 and 6.

(16) Scaling of H_I and account of pattern 1

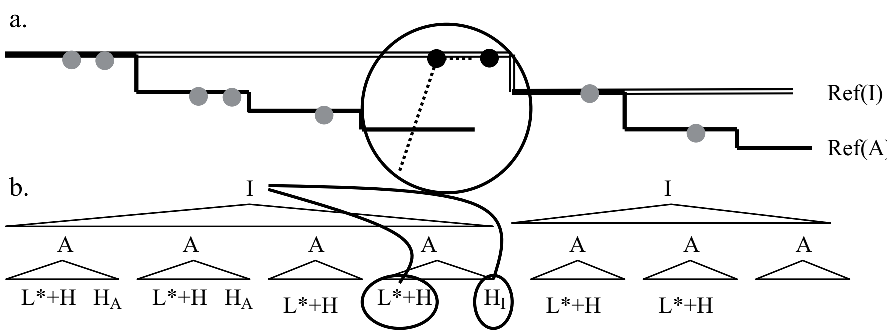


Pattern 2 is similar to pattern 1, except for the deletion of H_A and the possible boosting effect on the downstepped nuclear rise, as discussed in connection with Figs. 7 and 8.

To account for the difference between pattern 1 (and 2, though less clearly) on the one hand (downstepped nuclear rise) and patterns 3 and 4 on the other (upstepped nuclear rise), we need to postulate that the nuclear pitch accent allows two different associations. Since it marks the strongest stress of an A (the last A in I), it can be associated with A as in (14)–(16), in which case it is scaled to the downstepped register of A. At the same time, it marks the strongest stress of the intonation phrase, and can thus be associated with I as in (17). By (13), this leads to upstepped scaling of the reference-line of I as shown in (17). (17) is the account of pattern 3, with H_A deleted and both the nuclear rise and H_I upstepped; see the discussion of Figs. 9 and 10.

So far, the account captures the downstepped scaling of non-final pitch accents, the scaling of H_A, the upstepped scaling of H_I in patterns 1–3, and the options of scaling the nuclear pitch accent: downstepped in pattern 1 (and 2), where it can be construed as the strongest accent of an accentual phrase, and upstepped in pattern 3, where it can be construed as the strongest accent of an intonation phrase.

(17) Upstepped scaling of the nuclear rise and account of pattern 3

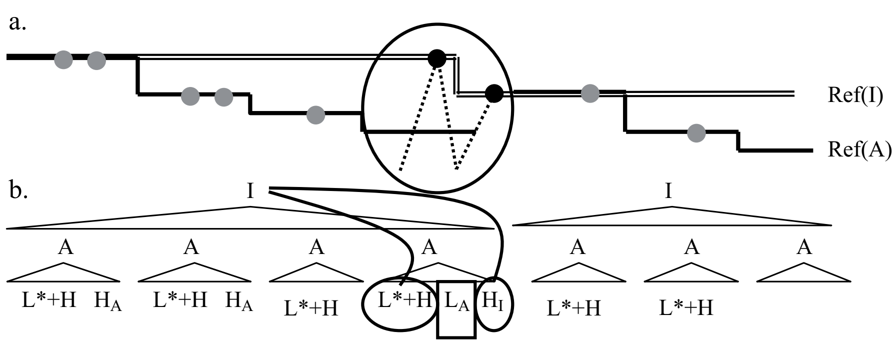


5.4. Pattern 4 and more on the position of lowering of the reference-lines

In pattern 4, the upstepped scaling of the nuclear accent can also be accounted for by association of the nuclear pitch accent with the intonation phrase. The scaling of H_I in pattern 4, however, is not obvious. Even though only the data of one speaker is at issue, this is discussed here in some detail, since it points in an interesting direction. The general account of scaling in (12) and (13) leads to the conclusion that H_I is scaled to the reference-line of I, which is reasonable, given that H_I is in fact scaled to the height of the phrasal reference-line in patterns 1–3. Observe now that H_I in Fig. 13, labeled 'H%' there, is dramatically lower than the preceding nuclear upstepped values (a Paired sample t-test confirms that this distance is significant; $t = 11.7[df=71]$, $p<.001$), and so cannot plausibly be scaled on a reference-line of I that continues at the same height after the nuclear upstepped values. At the same time, H_I is close in height to the peak in partial reset (H' in Fig. 13). Since there are reasons to believe that H_I is scaled to the reference-line of I, its scaling in pattern 4 would appear to suggest that the reference-line of I is not literally lowered between the two intonation phrases, but (in pattern 4) between the upstepped nuclear rise and H_I, as shown in (18).

A possible reason for early lowering of the reference-line is readily found: the presence of L_A preceding H_I. A L-tone is crucial for the application of downstep in many tone languages; see Connell and Ladd (1990) and Laniran and Clements (2003) for Yoruba. Also, in the utterance-final sequence L*+H H+L* in Fig. 2 there is arguably no downstep of the final pitch accent (cf. Truckenbrodt 2004). This can also be seen in Fig. 4, where the final plateau for speakers AT, BK, NA, MG, and TL is L*+H H+L*, and the height of the values clearly suggests absence of downstep within this final plateau. On the other hand, there is downstep between any two L*+H L*+H pitch accents (unless upstep/partial reset comes into play). For speaker FS, who uses L*+H as the utterance-final pitch accent, downstep also extends to the final L*+H, as can be seen in Figs. 4 and 6. A plausible source of the difference between the downstep in L*+H L*+H and the absence of downstep in L*+H H+L* is the presence of L between the H tones in the former case but not in the latter.

(18) Account of pattern 4: early lowering of the reference-line of I



If the reference-line of I lowers with L_A before the edge of I as in (18), the lowering cannot literally be triggered by the prosodic edges of I. There are independent indications that the prosodic trigger of downstep is not located in the prosodic edge, but in the position of the prominence corresponding to a prosodic constituent (see section 6.3. below). The revision of (10a) in (19a), retaining (10b) in (19b), takes the strongest stress of a particular prosodic level to be the trigger of lowering of the reference-line of that level, and thus allows the addition that the lowering is actually executed with a L tone that follows the stress.

(19) **Prosodic constituents and reference-lines** (revised from (10))

a. To each prosodic level X, assign a reference-line Ref(X) for H tones; lowering of this reference-line is triggered by the strongest stress of a constituent of level X, call it S(X); this lowering takes place with the first L tone that follows the starred tone on S(X).
b. For a prosodic constituent X which is initial in a higher prosodic constituent Y, Ref(X) = Ref(Y).

The effects of lowering on the level of the accentual phrase are shown in (20). A circle highlights the position of stress that triggers the lowering of a reference-line (here: by highlighting the starred tone on it). The stepped-down reference-line shows the position of lowering in the position of the next following L tone (boldfaced).

(20) Reference-line of A

a.

$)_A($

L_1*+H_1 H_A **L_2*** + H_2 downstepping L*+H accents

b.

$)_A($

L_1*+H_1 H_A H_2 + **L_2*** utterance-final plateau

In (20a), a sequence of downstepping L*+H accents, the lowering of the reference-line of A is triggered by the A-stress on L_1*+H_1, but the lowering takes effect with the next L tone that follows L_1*, which is L_2*. Correctly, this lowering does not affect H_A, which is scaled to the same height as H_1, but does affect H_2 on the next pitch accent. In the utterance-final (20b), the stress of L_1*+H_1 would trigger downstep with the next L following L_1*. This L tone, however, is the final L_2*, so the lowering of the reference-line does not affect H_2. This correctly predicts that H_2 is not downstepped relative to H_1 in (20b).

The lowering of the reference-line of I is shown in (21).

(21) Reference-line of I

a.

$]_I[$

L_n*+H_n (H_A) H_I **L_i***+H_i patterns 1−3

b.

$]_I[$

L_n*+H_n **L_A** H_I L_i*+H_i pattern 4

Patterns 1–3 have the tonal sequence (21a). The nuclear (strongest) stress of I triggers lowering of the reference-line of I by (19a), which occurs with the first

L after the pitch accent on the nuclear stress. This crucial L tone is here L_i. Thus, in patterns 1–3 the reference-line remains horizontal throughout the first I: upstepped H_I thus targets a reference-line of I of not-yet-lowered height, as desired. The lowering of the reference-line at point L_i* then correctly affects H_i, the peak in partial reset. Consider then pattern 4 in (21b). Here L_A is the first L tone after the nuclear pitch accent, and thus the position of lowering of the reference-line of I by (19a). This correctly allows the nuclear rise to be fully upstepped, and correctly predicts that H_I occurs at a downstepped height, comparable to the height of the following peak in partial reset.

With an account of the scaling of H_I in pattern 4, we can now also account for the fact that H_I is minimally higher than the H in partial reset (see Fig. 13). Both are scaled to the downstepped reference-line of I; however, H_I (but not the H in partial reset) is plausibly subject to phonetic dissimilation relative to the preceding L_A of pattern 4. A coherent picture of dissimilatory effects for the sequence /H_A H_I/ is obtained:

(22)	*phonological dissimilation*	*surface*	*phonetic dissimilation*[8]
Pattern 1:	(none)	H_A H_I	H_A lowers a bit, H_I raises a bit
Patterns 2,3:	$H_A \rightarrow ø$ / __ H_I	H_I	none: no T_A left to dissimilate from
Pattern 4:	$H_A \rightarrow L_A$ / __ H_I	L_A H_I	H_I raises a bit

As throughout, phonetic dissimilation is taken to be overlaid on the downstepped or upstepped target-values on the reference-lines and to lead to small deviations from these target-values.

5.5. Summary of the account of scaling

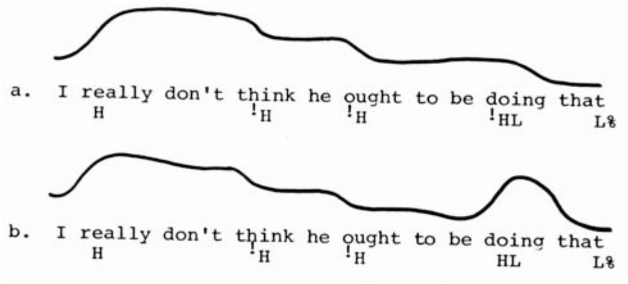
The German data provide evidence for a development of the model of Ladd (1988) and van den Berg, Gussenhoven, and Rietveld (1992) in such a way that prosodic constituents mediate the assignment of reference-lines to tones. Reference-lines of different prosodic levels are identified initially in higher prosodic constituents, and are lowered by stress on the relevant level. Tones are associated with prosodic constituents in simple and systematic ways, and this association goes a long way in predicting which reference-line a tone is scaled on. Edge tones are uniquely assigned the prosodic constituent of which they mark the edge, and are thus uniquely assigned to a phonetic reference-line. Pitch accents are associated with a prosodic constituent of which they mark the strongest stress; in the case of the nuclear pitch accent, this allows a choice among two levels of association. Association with the lower accentual phrase leads to scaling that reflects the downstep of earlier accentual phrases; association with the higher intonation phrase leads to scaling on the reference-line of the intonation phrase that is independent of this accentual downstep.

6. Connections to intonation in other languages

6.1. Upstep in utterance-final position

Ladd (1983) observes the two renditions of the same sentence in (23). Downstep affects all accents including the nuclear one in (23a), and all accents except for the nuclear one in (23b).

(23)



It is tempting to see the rendition in (23b) as upstepped by a return to a phrasal reference-line on the nuclear accent.

Consider also the suggestions about scaling the Swedish sentence accent by Bruce and Gårding (1978). Their illustration is reproduced in Fig. 14. Here the sentence accent is scaled in a return to a high reference-line for the dialects of Dalarna, Stockholm and Göteborg, after preceding lowering of words carrying Accent 2. Importantly, the high realization of the sentence accent is not a feature of utterance-medial position.

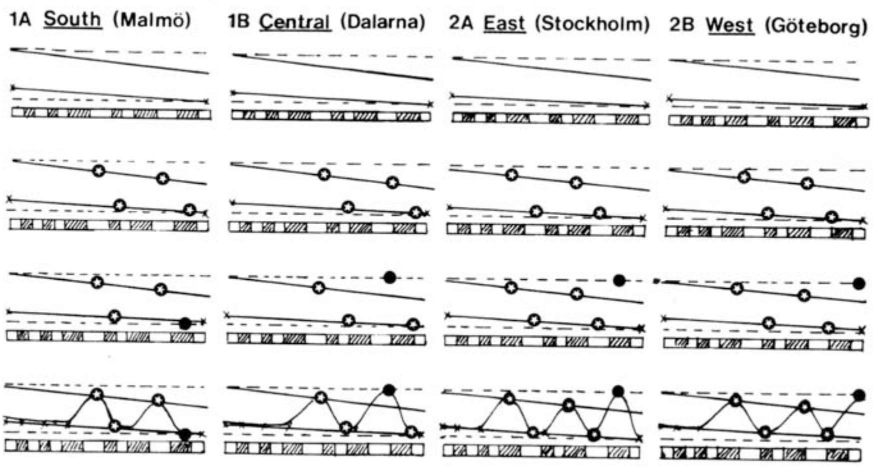


Figure 14. Phonetic derivations for the Swedish utterance *några långa nunnor*, [nɔra lɔŋːa nʉnːɔr], 'some tall nuns' in statement rendition in different dialects, illustrating suggestions of scaling by Bruce and Gårding (1978). The highs and lows for Accent 2 on the first syllables of [lɔŋːa] and of [nʉnːɔr] (starred dots) are introduced on the (solid) falling topline and baseline, respectively. The sentence accent (black dot) is introduced on the (dotted) higher and independent focal line for the dialects of Dalarna, Stockholm and Göteborg.

In our data, upstep is limited to positions preceding a medial I-boundary, and is not found in utterance-final position. Trivially, upstep on H_I is not found in utterance-final position in the recordings discussed here, since these recordings do not have any H_I tones in utterance-final position. However, the present account allows the choice of scaling a nuclear accent to the upstepped register of I, rather than to the downstepped register of I. Yet the speakers of patterns 3 and 4 who upstep on the nuclear accent in medial position do not upstep on the nuclear pitch accent in utterance-final position. This can be seen in the final pitch accents in Figs. 2, 3, 10, and 12. The observations on English and on Swedish suggest that it may not be right to generally restrict upstep to non-final position. Truckenbrodt (2002) offers the hypothesis that medial upstep is favored since it makes the lowering of the reference-line of I 'visible' in adjacent pitch accents – the upstepped pitch accent and the one in partial reset. That hypothesis is compatible with our results, on the understanding that such a motivation is a necessary, but not a sufficient condition for association of a nuclear pitch accent with the intonation phrase in German. Thus, this motivation would be present in medial position, allowing the speakers of patterns 3 and 4 to choose nuclear association with the intonation phrase. At the same time, the presence of such motivation would still have to allow the choice of a downstepped medial pitch accent, as was found in patterns 1 and 2. In final position, in the absence of such motivation for higher association and scaling, upstep could then correctly be predicted to be absent in German for all speakers. Further, by the account above, H_I is necessarily scaled to the reference-line of the intonation phrase in German.

Different motivations for the high scaling in non-final positions (23) and Figure 4 could then be hypothesized. In the English example (23), such a phonological choice might be related to a preference for high scaling in emphatic speech. In the Swedish cases in Figure 14, it would follow if the sentence accent is an edge tone of a higher prosodic constituent. In Riad (1998), the Prominence tone (the sentence accent, for example, of the Stockholm contour) is right-aligned with the ‘Focus phrase’. The alignment-behavior may be taken to indicate that this is indeed an edge tone. It also seeks out a prominent syllable for association; this may be taken to be an additional property of certain edge tones, as argued in Gussenhoven (2000a,b) for Roermond Dutch.

6.2. The OCP in intonation

The OCP disallows adjacent identical elements such as adjacent identical H tones, and has originally been argued for in tone languages; see Myers (1997) and Yip (2007) for recent reviews. Hayes and Lahiri (1991) argue that in the intonation of Bengali, the OCP is active at the level of the tune, such that in a sequence of tones in an intonation phrase, two H tones must always be separated by L. Since English allows H tones that are adjacent on the tonal tier, but does not allow two H tones in a bitonal pitch accent according to Beckman and Pierrrehumbert (1986), Hayes and Lahiri point out that English may obey the OCP in the more narrow domain of the pitch accent. In the account of the German data above, adjacent H tones are also allowed, for example in L*+H H_A and in L*+H H+L*. As in English, however, bitonal pitch accents with two like tones seem not to exist in German (Grice and Baumann 2002 and Grice, Baumann, and Benzmüller 2005). In addition to the tones of a bitonal pitch accent, the sequence /H_A H_I/ seems to be affected by the OCP for the speakers of patterns 2–4, where H_A is deleted or dissimilated to L when preceding H_I (see the summary in (22)). One might say, then, that the speaker of pattern 1 employs the OCP within the pitch accent only, as do speakers of English, while the speakers of patterns 2–4 employ a domain of application of the OCP that is intermediate between the domain of the OCP in English and in Bengali. One may characterize this domain by building on Hayes and Lahiri's suggestion that the pitch accent is in some sense a domain. If we write the 'edges' of such domains as '|', a representative sequence of tones from the German material would be divided as in (24).

(24) [L*+H | H_A | L*+H | H_A H_I] [(...) | L*+H | H_A | H+L* | L_I]

It can be seen that adjacent H tones are always separated by the 'edge' of a pitch accent, except in the case of the sequence /H_A H_I/. A possible generalization for the speakers of pattern 2–4 is then that, for these speakers, the OCP is bounded by ‘edges’ of pitch accents. This would correctly allow [L*+H H_A] with an ‘edge’ of a pitch accent intervening between the two H tones, and it would correctly disallow /H_A H_I/ with no such intervening ‘edge’ of a pitch accent.

6.3. Prominence and downstep

The suggestion in (19a) that downstep is linked to *stress* is preliminary. If, however, downstep is in general an aspect of enhancing the *prominence* of an earlier element, the crucial feature of downstep lies in increasing the distance between the enhanced element and following material (Truckenbrodt 2004). There are a number of indications in the literature that point in this direction.

In Tokyo Japanese (Poser 1984; Pierrehumbert and Beckman 1988; Kubozono 1989), accentual phrases can be accented (carry an H*L accent) or unaccented (not carry such an accent). Downstep does not affect all accentual phrases (which would be parallel to (10a)). Rather, it is observed only following accents (in parallel to the revised (19a)). Further, Poser (1984) argues that downstep by an HL accent takes effect with the L tone of the accent, and Pierrehumbert and Beckman 1988 note that this is compatible with their data. At the same time, the H tone of the accent in Tokyo Japanese shows unexpected height in comparison with H tones that are not part of an accent, all else being equal. These observations are compatible with the interpretation that the prominence of the accent is enhanced by raising the H of the accent and by downstep of following elements, beginning with the accentual L.

In the tone language Mandarin Chinese (Xu 1999), focus has the simultaneous effects of lowering and compressing the range of following tones, while raising the range inside of the focus. (This process is not dependent on the presence of L tone.)

In the tone language Yoruba, Laniran and Clements (2003) have shown that two partly overlapping phenomena are each real in their own right, namely downstep in the sequence HLH and raising of H before L. Yet the two phenomena seem to go hand in hand. For example, both processes are simultaneously blocked by intervening M(id) tone in the sequence [H M L M H]. This is compatible with an interpretation in which the prominence of an H tone, in the presence of following L, is enhanced by raising its own height and by concomitant compression of the following range by downstep.

Finally, Truckenbrodt (2004) argues, on the basis of the German data that the present paper also draw on, that H tones preceding accentual downstep are realized slightly but significantly higher than H tones not preceding accentual downstep, all else being equal. This also supports an interpretation by which downstep is linked to the enhancement of prominence of an earlier element. In the German case, the preceding prominence thus enhanced would be the stress of an earlier element, in line with the formulation in (19a).

The intonation of the speaker of pattern 4, which led to the formulation in (19a), thus suggests that prominence can lower the following range (i.e. lower a reference-line) at different prosodic levels, with stronger stress (the strongest of I) lowering the reference-line of I, and lesser stress (the strongest of A) lowering the reference-line of A.

7. Summary

The German data support a development of the models of Ladd (1988) and van den Berg, Gussenhoven, and Rietveld (1992) in terms of the assignment of tones to reference-lines on the basis of prosodic constituents, as formulated in (12), (13), and (19). Tones pertaining to accentual phrases are scaled to the reference-line that is lowered once per accentual phrase. Tones pertaining to intonation phrases are scaled to the (phrasal) reference-line that is lowered once per intonation phrase, i.e. they are upstepped in the normal case. The data point towards a more general account in which lowering of a reference-line is triggered by the stress on the relevant prosodic level, and is executed with the next L tone following the starred tone on that stress. This is compatible with a perspective in which lowering and compression of the tonal range is more generally an effect of the enhancement of the prominence of an earlier element.

Four different contours preceding the medial I-boundary (patterns 1–4) were observed. They can be reduced to a single underlying specification L*+H H_A H_I. Individual variation is allowed in regard to the association of the nuclear L*+H with A (resulting in downstepped scaling) or with I (resulting in upstepped scaling). Other individual differences can be reduced to individual variation in dissimilatory effects in the sequence H_A H_I, as summed up in (22). The phonological effects of (22) can plausibly be related to the OCP.

Acknowledgements

I would like to thank audiences at the *International Conference on Tone and Intonation* (2004 in Santorini) and at Amherst University for useful questions and comments. Thanks also to Lisa Selkirk and Bob Ladd for helpful advice in connection with earlier versions of this work and to an anonymous reviewer for helpful comments on the present paper. All errors are of course my own. This work benefited from financial support through the DFG-project B15 of the SFB 441 in Tübingen.

Notes

1. See van den Berg, Gussenhoven and Rietveld (1992) for empirical confirmation of this property of the partial reset.
2. The term *upstep* is also used for phenomena in other languages to which this analysis does not apply, such as for higher than expected scaling of L% and H% English edge tones following an H– phrase accent since Pierrehumbert (1980), and for the successive raising of the tonal height in Kimatuumbi by Odden (1995).
3. See Uhmann (1991) for detailed suggestions about the application of the SAAR to German; see Truckenbrodt (2002, 2004) for details of this application in the data discussed in the present paper. The term *accentual phrase* (Beckman and Pierrehumbert 1986, Pierrehumbert and Beckman 1988) is used here for the domains of accent assignment by the SAAR.
4. See for example Selkirk (1984), von Stechow and Uhmann (1986), Cinque (1993); see Truckenbrodt (2006) for discussion of the assignment of phrasal stress.
5. Speaker SW was recorded separately before the other speakers, with stimuli that had only two accents in the second clause of set a.
6. Alternatives to the deletion analysis are not altogether excluded here. However, the deletion analysis not only captures a systematic distinction in the course of F_0 between pattern 1 and pattern 2 as described, but also allows to connect this to the height of H_I above that of the utterance-initial rise in pattern 1 (dissimilation from H_A) in contrast to the absence of such an effect of additional height in pattern 2 (no dissimilation from H_A due to deletion of H_A). There is further evidence for phonological and phonetic dissimilation of H_A and H_I from each other (summed up in (22) below), which indirectly supports this analysis, and also provides a plausible motivation for deletion of H_A before H_I.
7. See Liberman and Pierrehumbert (1984) on the special status of the utterance-final L edge tone of the intonation phrase in English. It seems to be scaled to the bottom of the speaker's range. This understanding of the scaling of the final L is compatible with the German data reported here. Therefore L_I is used as a reasonable reference-point.
8. An alternative account of the raised H_I values might pursue the notion that the preceding T_A in some sense, 'pushes H_I outside' the edge of I and into the following breath-pause phonetically. Thus, the unexpectedly high values of H_I after T_A (patterns 1 and 4) were typically rising steeply right up to the following breath-pause, with no onset of a fall before the breath-pause. This was typically not the case for H_I values of the other speakers,

(patterns 2 and 3), where H_I is not preceded by T_A and does not show an effect of phonetic raising relative to the expected values.

References

Atterer, Michaela, and D. Robert Ladd
2004 On the phonetics and phonology of "segmental anchoring" of F_0: Evidence from German. *Journal of Phonetics* 32: 177–197.

Barker, Geoffrey S.
2002 Intonation patterns in Tyrolean German: An autosegmental-metrical analysis. Ph. D. diss., University of California, Berkeley.

Beckman, Mary E., and Janet B. Pierrehumbert
1986 Intonational structure in Japanese and English. *Phonology Yearbook* 3: 255–309.

Bruce, Gösta, and Eva Gårding
1978 A prosodic typology of Swedish dialects. In *Nordic Prosody,* Eva Gårding, Gösta Bruce, and Robert Bannert (eds.), 219–228. Lund: Lund University.

Cinque, Guglielmo
1993 A null theory of phrase and compound stress. *Linguistic Inquiry* 24: 239–297.

Connell, Bruce, and D. Robert Ladd
1990 Aspects of pitch realisation in Yoruba. *Phonology* 7: 1–29.

Féry, Caroline
1993 *German Intonational Patterns*. Tübingen: Niemeyer.

Féry, Caroline, and Hubert Truckenbrodt
2005 Sisterhood and tonal scaling. *Studia Linguistica* 59, 223–243.

Fitzpatrick-Cole, Jennifer
1999 The Alpine intonation of Bern Swiss German. *Proceedings of the 14th International Congress of Phonetic Sciences*, John J. Ohala (ed.), 941–944. San Francisco.

Gilles, Peter
2005 *Regionale Prosodie im Deutschen. Variabilität in der Intonation von Abschluss und Weiterweisung*. Berlin/New York: de Gruyter.

Grabe, Esther
1998 Comparative intonational phonology: English and German. Ph. D. diss., Radboud University Nijmegen.

Grice, Martine, and Stefan Baumann
2002 Deutsche Intonation und GToBI. *Linguistische Berichte* 191: 267–298.

Grice, Martine, Stefan Baumann, and Ralf Benzmuller
2005 German intonation in autosegmental-metrical phonology. In *Prosodic Typology: The Phonology of Intonation and Phrasing,* Sun-Ah Jun (ed.), 55–83. Oxford: Oxford University Press.

Gussenhoven, Carlos
1983 Focus, mode and the nucleus. *Journal of Linguistics* 19: 377–417.

1992 Sentence accents and argument structure. In *Themantic Structure: Its Role in Grammar,* Ignacio M. Roca (ed.), 79–106. Berlin/New York: Foris.

2000a The boundary tones are coming: On the nonperipheral realization of boundary tones. In *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, Michael B. Broe, and Janet B. Pierrehumbert (eds.), 132–151. Cambridge: Cambridge University Press.

2000b The lexical tone contrast of Roermond Dutch in Optimality Theory. In *Prosody: Theory and Experiment,* Merle Horne (ed.), 129–167. Dordrecht: Kluwer.

2004 *The Phonology of Tone and Intonation*. Cambridge: Cambridge University Press.

Gussenhoven, Carlos and Toni Rietveld

1988 Fundamental frequency declination in Dutch: testing three hypotheses. *Journal of Phonetics* 16: 355–369.

Hayes, Bruce, and Aditi Lahiri

1991 Bengali intonational phonology. *Natural Language & Linguistic Theory* 9: 47–96.

Kubozono, Haruo

1989 Syntactic and rhythmic effects on downstep in Japanese. *Phonology* 6: 39–67.

Kügler, Frank

2005 Swabian and Upper Saxon intonational patterns. Ph. D. diss., Universität Potsdam.

Ladd, D. Robert

1983 Phonological features of intonational peaks. *Language* 59: 721–759.

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

1993 On the theoretical status of 'the baseline' in modeling intonation. *Language and Speech* 36: 435–451.

1996 *Intonational Phonology*. Cambridge: Cambridge University Press.

Laniran, Yetunde O., and G. Nick Clements

2003 Downstep and high raising: interacting factors in Yoruba tone production. *Journal of Phonetics* 31: 203–250.

Liberman, Mark, and Janet B. Pierrehumbert

1984 Intonational invariance under changes in pitch range and length. In *Language, Sound, Structure: Studies in Phonology Presented to Morris Halle by his Teacher and Students,* Mark Aronoff, and Richard T. Oehrle (eds.), 157–233. Cambridge: MIT Press.

Myers, Scott

1997 OCP effects in Optimality Theory. *Natural Language and Linguistic Theory* 15: 847–892.

Nespor, Marina, and Irene Vogel

1986 *Prosodic Phonology*. Dordrecht: Foris.

Odden, David

1995 Tone: African languages. In *The Handbook of Phonological Theory,* John Goldsmith (ed.), 444–475. Blackwell.

Peters, Jörg
2002 Intonation und Fokus im Hamburgischen. *Zeitschrift für Sprachwissenschaft* 189: 27–57.
Pierrehumbert, Janet B.
1980 The phonology and phonetics of English intonation. PhD dissertation, MIT.
Pierrehumbert, Janet B., and Mary E. Beckman
1988 *Japanese Tone Structure*. Cambridge, Mass.: MIT Press.
Poser, William J.
1984 The phonetics and phonology of tone and intonation in Japanese. Ph. D. diss., MIT.
Prieto, Pilar, Chilin Shih, and Holly Nibert
1996 Pitch downtrends in Spanish. *Journal of Phonetics* 24: 445–473.
Riad, Tomas
1998 Towards a Scandinavian accent typology. In *Phonology and Morphology of the Germanic Languages,* Wolfgang Kehrein, and Richard Wiese (eds.), 77–109. Tübingen: Niemeyer.
Selkirk, Elisabeth
1984 *Phonology and Syntax: the Relation between Sound and Structure*. Cambridge, Mass: MIT Press.
Shih, Chilin
2000 A declination model of Mandarin Chinese. In *Intonation. Analysis, modeling and technology,* A. Botinis (ed.), 243–268. Dordrecht: Kluwer.
Truckenbrodt, Hubert
2002 Upstep and embedded register levels. *Phonology* 19: 77–120.
2004 Final lowering in non-final position. *Journal of Phonetics* 32: 313–348.
2006 Phrasal Stress. In *The Encyclopedia of Languages and Linguistics, 2nd edition*, *Vol. 9*, K. Brown (ed.), 572. Oxford: Elsevier.
Truckenbrodt, Hubert, and Caroline Féry
2005 More on hierarchical structure and tonal scaling. Manuscript. University of Tübingen and University of Potsdam.
Uhmann, Susanne
1991 *Fokusphonologie*. Tübingen: Niemeyer.
Van den Berg, Rob, Carlos Gussenhoven, and Toni Rietveld
1992 Downstep in Dutch: Implications for a model. In *Papers in Laboratory Phonology II: Gesture, Segment, Prosody,* G. J. Docherty, and D. R. Ladd (eds.), 335–367. Cambridge: Cambridge University Press.
Von Stechow, Arnim, and Susanne Uhmann
1986 Some remarks on focus projection. In *Topic, Focus, and Configurationality*, Werner Abraham and Sjaak de Meij (eds.), 295–320. Amsterdam: John Benjamins.
Xu, Yi
1999 Effects of tone and focus on the formation and alignment of F_0 contours. *Journal of Phonetics* 27: 55–105.
Yip, Moira
2007 *Tone*. In *The Cambridge Handbook of Phonology*. Paul de Lacy (ed.), 229–251. Cambridge: Cambridge University Press.