Phonology **19** (2002) 77–120.
DOI: 10.1017/S095267570200427X Printed in the United Kingdom

*Upstep and embedded register levels**

Hubert Truckenbrodt
University of Tübingen

This article is concerned with an interesting upstep phenomenon in the intonation of some speakers of Southern dialects of German. It experimentally establishes the main properties of this upstep phenomenon, and discusses the theoretical consequences. The upstep occurs on the nuclear pitch accent of a non-final intonation phrase. It targets the phonetic height of the utterance-initial peak, regardless of downstepped peaks that intervene between the initial peak and the upstepped peak. The findings are argued to provide unexpected support for a model of intonation in which downstep among accents can be embedded inside downstep among larger prosodic domains (Ladd 1988, van den Berg *et al*. 1992). In a combination of that model with an extension of Pierrehumbert & Beckman (1988), it is suggested that the choice between downstep and upstep is conditioned by association to higher prosodic constituents in a systematic way.

In this article, I present an experimental investigation of an unexpected upstep phenomenon found in the intonation of some (though not all) speakers from Southern Germany. The upstep phenomenon occurs on the nuclear pitch accent of a non-final intonation phrase. Following downstep on prenuclear pitch accents, the nuclear pitch accent is scaled to a height comparable to the utterance-initial peak, undoing preceding downstep. The upstep is followed by an initial peak in the following clause that is downstepped relative to the upstepped peak, and at the same time partly reset relative to earlier downstep. I will argue that the phenomenon provides a new kind of support for the analysis of partial reset by Ladd (1988) as targeting an abstractly downstepped register level, and further

* I would like to thank Bob Ladd, Carlos Gussenhoven and an anonymous reviewer of *Phonology* for many helpful comments and suggestions in connection with this article. For comments on earlier versions of this work, thanks also to Akin Akinlabi, Paul de Lacy, Caroline Féry, Jennifer Fitzpatrick, Bruce Hayes, Michael Kenstowicz, Aditi Lahiri, Alan Prince, Lisa Selkirk, Donca Steriade and Moira Yip, as well as audiences at Constance University, McGill University, MIT, Rutgers University, the University of Edinburgh, the University of Maryland, the University of Potsdam, the Workshop on Focus, Intonation and Phrasing in Freudental 1998, the Workshop *Interaktion der Grammatikmodule* at the DGfS meeting 1997 and the *Zentrum für Allgemeine Sprachwissenschaften*, Berlin. Many thanks also to Judy Bauer and the editors of *Phonology* for valuable help with the presentation of the final version, and to Susi Wurmbrand for volunteering her voice for many recordings leading up to the present work. All errors are of course my own. This work was supported by Rutgers University Research Council Grants #2-02189 and #2-02284.

supports the phonetic implementation of this idea in van den Berg *et al.* (1992). I account for the phenomenon by combining the phonetic model of van den Berg *et al.* (1992) with proposals about phonological association from Pierrehumbert & Beckman (1988), and an extension of the suggestion of Pierrehumbert & Beckman (1988) about the prosodic conditioning of register scaling.

The article is structured as follows. In §1, I introduce some background on the phonology and phonetics of the approach to intonation adopted here, illustrating this with a German example. In §2, I introduce the phenomenon of upstep and its phonetic analysis. In §3, the experimental evidence for the phonetic analysis is presented. The phonological part of the analysis is discussed in §4. §5 sums up the results.

1 Background on intonational phonology and phonetics

The last quarter of a century has seen impressive developments in the theory of intonation, on which the present work builds. Pierrehumbert (1980) integrated suggestions from Bruce's (1977) and Bruce & Gårding's (1978) work on Swedish into a framework for intonational phonology with an exploration of English intonation. A corresponding theory of phonetic implementation, originating in Pierrehumbert (1980), was further developed in Liberman & Pierrehumbert (1984). Pierrehumbert & Hirschberg (1990) further extended the theory of intonation to account for the pragmatic meanings of the parts of intonation contours. Originally developed for English, this theory was extended to Japanese in Pierrehumbert & Beckman (1988), with a comparison of English and Japanese offered in Beckman & Pierrehumbert (1986), and has since been applied to many other languages; see Ladd (1996) for an overview. Modifications of this theory that have also had some influence include Gussenhoven (1984) and Ladd (1983a, 1988, 1990). Recent topics of investigation include, on the phonological side, the application of Optimality Theory (Prince & Smolensky 1993) to the placement of tones in intonation (Pierrehumbert 1994, de Lacy 1998, Gussenhoven 2000b) and the unexpected phenomenon of edge tones seeking stress (Ladd 1996, Grice *et al.* 2000, Gussenhoven 2000a), and on the phonetic side, the timing of tones relative to segmental and syllabic structure (Silverman & Pierrehumbert 1990, Prieto *et al.* 1995, Xu 1998, Arvaniti *et al.* 2000, Ladd *et al.* 2000; see also Myers 1999) and investigations of the perception of peak prominence and/or declination (Terken 1991, 1993, Terken & Hermes 2000, Ladd 1993a, Ladd *et al.* 1994, Ladd & Morton 1997, Gussenhoven *et al.* 1997, Gussenhoven & Rietveld 1998). This approach has been applied to German intonation by Wunderlich (1998), Uhmann (1991), Féry (1992, 1993), Büring (1995), Grabe (1998), Fitzpatrick-Cole (1999) and Grice *et al.* (to appear).[1]

[1] For other discussions of German intonation, see von Essen (1964), Isačenko & Schädlich (1966), Bierwisch (1966), Pheby (1984) and the contributions in Altmann (1988), Altmann *et al.* (1989) and Kohler (1991).

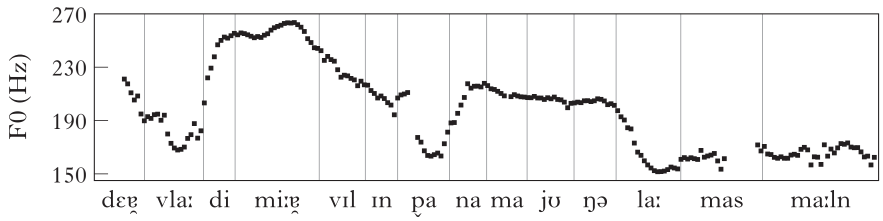


Figure 1

F0 track for *Der Wladimir will in Panama junge Lamas malen*. 'Vladimir wants to paint young llamas in Panama.' Speaker SW. Labelling is by syllables, as far as possible.

The present work concentrates on an aspect of German intonation that has not received much attention in previous work. However, I have found brief discussions of a phenomenon in English which appears to be related in Ladd (1983a) and in Beckman & Pierrehumbert (1986), and similarly brief discussions of the phenomenon in German in Féry (1993) and Fitzpatrick-Cole (1999). I return to these below. This phenomenon, in the form in which it is investigated here, only emerges with utterances of a certain complexity in German. The theoretical implications of this phenomenon depend on the analysis of simpler and more standard phenomena in the intonational system. In this section, I therefore introduce some basic concepts of the analysis of intonation adopted here, and illustrate these concepts with a German example. The basic concepts introduced here, as well as their applications to German, provide relevant background to the discussion of the more complex intonational patterns in German in the remainder of this article.

1.1 Phonology: tones in intonation and prosodic structure

Consider the F0 track in Fig. 1, from my German material. This is the principal acoustic correlate of what listeners perceive as the sentence melody. To the extent that it is representative (see below), it is also what the theory of the phonology and the phonology–phonetics mapping in intonation seeks to approximate in its predictions for a phonetic output.[2]

In the tone-sequence analysis of F0 in intonation adopted here, the analysis of such contours postulates a sequence of phonological H or L tones. Each phonological tone is taken to define a phonetic point. The predicted phonetic output then results from linear interpolation between

[2] Labelling is by syntactic word boundary in [vɪl ɪn]. [ŋ] in [jʊŋə] is ambisyllabic on standard assumptions on German syllable structure (Féry 1995, Vennemann 1982, Wiese 1996 and others), but labelled as part of the following syllable.

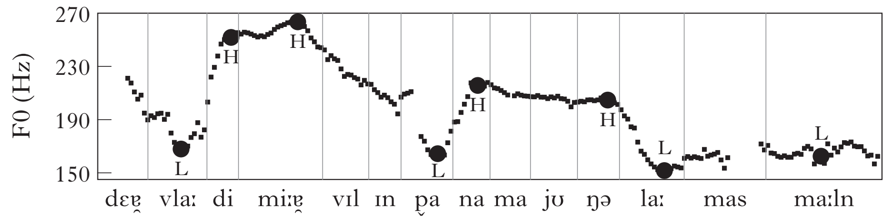


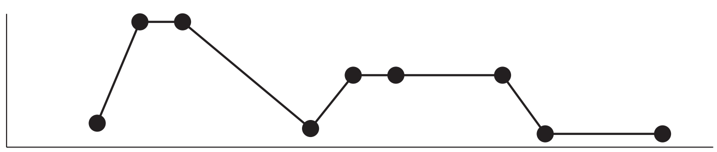
Figure 2

F0 track from Fig. 1, annotated with circles that approximate the turning points. The analysis postulates L and H tones, such that the turning points are the phonetic values of the L and H tones. The contour is taken to result from interpolation between these phonetic values.

the phonetic values of the tones, as shown in Fig. 2. In other words, each significant turning point in the F0 plot can be taken as evidence of the presence of a tone that defines this turning point.[3]

In the phonological analysis, H and L tones may either be assigned as pitch accents or as edge tones. Pitch accents are tones or sequences of tones associated with stressed syllables. They may be simple, such as L* or H*, or binary, such as L*+H. Following the notation in Pierrehumbert (1980), the star diacritic marks the tone that is associated with the prominent syllable in a pitch accent. The plus sign connects a starred tone to another tone with which it forms a binary pitch accent. As shown in (1), the present analysis postulates three pitch accents for this contour, two instances of L*+H in non-final position, and one instance of H+L* in final position. These are associated with the three beats of phrasal stress in an all-new rendition of this sentence as shown.

(1)



L*+H H_P L*+H H_P H+L* L%

(x $)_I$

(x) (x) (x $)_P$

dɛɐ̯ vla:.di.mi:ɐ̯ vɪl ɪn pa̯.na.ma jʊ.ŋə la:.mas ma:ln

Der Wladimir will in Panama junge Lamas malen.

'Vladimir wants to paint young llamas in Panama.'

Edge tones are associated with the edges of higher prosodic constituents. The notation here follows Pierrehumbert (1980) in using the

[3] Some languages, including English (Pierrehumbert 1980), may show sag instead of linear interpolation between H tones.

diacritic % for edge tones of the intonation phrase. For high edge tones of the phonological phrase (see immediately below) the subscript 'P' is adopted from Hayes & Lahiri (1991). With this, the analysis of the contour at hand postulates H_P edge tones at the right edge of non-final phonological phrases, and a L% edge tone at the right edge of the entire intonation phrase, as shown in (1).

The prosodic analysis in (1) follows Uhmann (1991)'s adaptation of Gussenhoven (1983) to German. The analysis distinguishes a lower prosodic level from a higher prosodic level. The lower level is called 'focus domain' by Gussenhoven (1983) and 'accent domain' by Uhmann (1991), following a suggestion by Ladd (1983b). I refer to it as P-PHRASE or P, which is shorthand for phonological phrase, following terminological choices and suggestions in Hayes & Lahiri (1991) and Truckenbrodt (1999). The higher prosodic level, the intonation phrase (*Intonations-phrase*) of Uhmann, is called I-PHRASE or I, adopting the shorthand for intonation(al) phrase from Hayes & Lahiri (1991). The strongest stress in the i-phrase is referred to as the NUCLEAR STRESS. Notice that, following the suggestion of Uhmann (1991) for German, the rightmost beat of stress in the i-phrase is strengthened in this example.[4] This is not immediately relevant to the case at hand, but will be important later in this article. For other issues in the assignment of prosodic structure in German, the reader is referred to §3.1.1 and the references there.

The tonal analysis is corroborated in two ways. First, the theory developed for other languages – incorporating findings in those languages – restricts the choices available. For example, each tone must be a part of a pitch accent or an edge tone. The three low points analysed as L* in this representation fall on the three syllables perceived as having the greatest prominence of this sentence, but none of these L tones is adjacent to a larger prosodic edge. They are thus plausible elements of pitch accents, but not plausible edge tones. Similarly, the H tones analysed as belonging to the pitch accents L*+H, L*+H and H+L* are on three syllables, none of which is adjacent to a larger prosodic edge. However, they are each adjacent to a stressed syllable, and so are plausibly analysed as elements of bitonal pitch accents. Conversely, the tones analysed as the edge tones H_P and L% in this analysis are adjacent to larger prosodic edges, but are not within plausible range of a larger beat of stress.

Second, the analysis is corroborated by repeated regular occurrences in similar contexts, which allows a regular tonal system to be inferred. For example, the speakers of my investigation come from the southern German-speaking area, including Austria. They all showed an almost exceptionless use of contours plausibly analysed as L*+H H_P in non-final position, but differed in their choice of final pitch accent, using H+L*, L*+H or L*. Each speaker, however, was internally consistent in her production. For speaker SW, from whom the F0 track in Figs. 1 and 2 is taken, contours that are plausibly analysed in terms of L*+H H_P in non-

[4] This is similarly suggested in Hayes & Lahiri (1991) for Bengali, and in Selkirk (1995) for English.

final position were found with almost no exceptions in literally hundreds of recordings of sentences produced as the answer to (the German analogue of) the question 'What's new?'. In final position, H+L* L% was regularly found, alternating sometimes with a downstep to a mid level on the syllable with nuclear stress (see below).

The analysis of a rise taking effect with the stressed syllable followed by a plateau, as L*+H followed by a H edge tone, parallels Beckman & Pierrehumbert (1986)'s analysis of a similar contour in English. The L*+H pitch accent has also been observed in non-final position in German declaratives by Uhmann (1991), Féry (1992, 1993), Grabe (1998) and Fitzpatrick-Cole (1999).[5] The plateau following L*+H was observed in German by Féry (1992). The present analysis is compatible with the transcription suggestions of German ToBI in Grice *et al.* (to appear), where the end of the plateau is analysed in terms of a H edge tone, as in Beckman & Pierrehumbert (1986)'s suggestion for English. The analysis of the final contour of SW in Fig. 2 as H+L* is also compatible with the suggestions of Grice *et al.* (to appear) for transcribing nuclear contours in German. The occasional alternation with a downstep to mid on the nuclear syllable might be the H+!H* of Grice *et al.* (to appear) or the !H*+L of Grabe (1998), who suggests that the latter may not be categorically distinct from H+L* in German. It can be seen in final position in the illustration in Fig. 3 below, where it is non-committally transcribed as !H*.

1.2 Phonetics: register in intonation

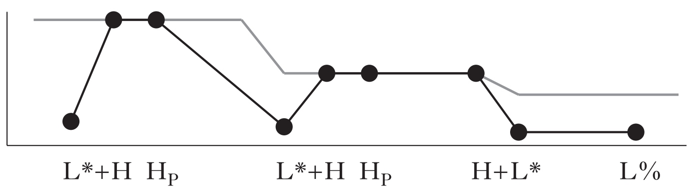
Let us then turn to more detail in the phonetic implementation of tones. It is apparent from Fig. 2 that not all tones of the same phonological polarity are of the same phonetic height. In German, as in many other languages, the phonetic differences are particularly striking among H tones, and I will limit my attention to these in the present article. A much-discussed phenomenon that leads to different phonetic height among H tones is that of downstep (Pierrehumbert 1980, Liberman & Pierrehumbert 1984). Under language-specific conditions, a H tone may be lower than a preceding H tone. The language-specific conditions may involve the presence of intervening L tone (Odden 1995) and may involve

[5] Uhmann (1991) shows the use of L*+H on a single prenuclear accent. In Féry (1992, 1993) and Büring (1995), it is discussed in connection with prenuclear (pragmatic) topics. Grabe (1998) successfully elicits sequences of prenuclear L*+H pitch accent with the help of syntactically coordinated lists. In the more Northern varieties of German discussed by these authors, the prenuclear L*+H is in competition with an also possible (prenuclear or nuclear) H*+L (or H* L-; see Grice *et al.* 2000 and references there). I believe that H*(+L) is restricted to contrastive contexts in the Southern German-speaking area, where the speakers reported on below come from, so that prenuclear L*+H could be elicited with regularity in non-final position, though the elements carrying them were neither topics in the relevant sense nor elements of lists. The observation in Fitzpatrick-Cole (1999) that L*+H is the default pitch accent in Berne Swiss German (including prenuclear position) is compatible with these remarks.

the presence of prominence in certain ways (Beckman & Pierrehumbert 1986, Pierrehumbert & Beckman 1988 for Japanese and English). The lowering of the second plateau relative to the first in Fig. 2 is analysed here as downstep (motivation for this analysis will be discussed shortly).

Since Clements (1979), downstep has often been analysed in terms of an abstract phonetic register. Register is usually construed as an interval on the vertical F0 scale, with H tones fixed to the top of the interval, and L tones fixed to the bottom of the interval. Downstep is then analysed as lowering and/or narrowing of the register interval, entailing lowering of the H tones fixed to the top of the register. Since I am concerned here only with the phonetic values of H tones, I will limit discussion to the top of the register. Downstep in Fig. 2 in terms of lowering of the top of a register is then analysed as shown in (2).

(2)



Downstep in German was observed by Féry (1993: ch. 5), and discussed in some detail in Grabe (1998: ch. 6). In both these works, it is discussed for tonal sequences analysed as H*+L. Fitzpatrick-Cole (1999) observes downstep in Berne Swiss German between a prenuclear and a nuclear L*+H. To see why the downstep analysis is plausible for SW and the other speakers in my sample, let us consider two other standard factors that have been observed to affect the phonetic height of H tones. DECLINATION, according to Pierrehumbert (1980) and Pierrehumbert & Beckman (1988), is a global gradual effect that can make values later in the utterance appear lower in absolute terms than values earlier in the utterance, all else being equal. In a metaphor of Pierrehumbert (1980), declination can be seen as a downward tilt and narrowing on graph-paper, on which the values predicted by the phonology–phonetics mapping are plotted.[6] One reason for not analysing the lowering of the tonal values in Fig. 2 as declination will become apparent below. Following Pierrehumbert (1980) and Pierrehumbert & Beckman (1988), declination

[6] Liberman & Pierrehumbert (1984) reanalyse the evidence for declination in Pierrehumbert (1980) as final lowering (see the next paragraph in the text). Gussenhoven & Rietveld (1988) argue that declination and final lowering both play a role simultaneously in the scaling of tones, and Pierrehumbert & Beckman (1988) also find separate effects of declination and final lowering in their Japanese data. In Pierrehumbert (1980), the downward tilt in declination is calculated against a falling reference line on which L% edge tones lie, but not other L tones in the utterance. The abstract nature of this reference line was later confirmed in the reinterpretation of the data of Terken (1991) by Ladd (1993a), accepted in the relevant respects by Terken (1993). See also Ladd *et al*. (1994), Gussenhoven *et al*. (1997), Ladd & Morton (1997), Gussenhoven & Rietveld (1998), Terken & Hermes (2000).

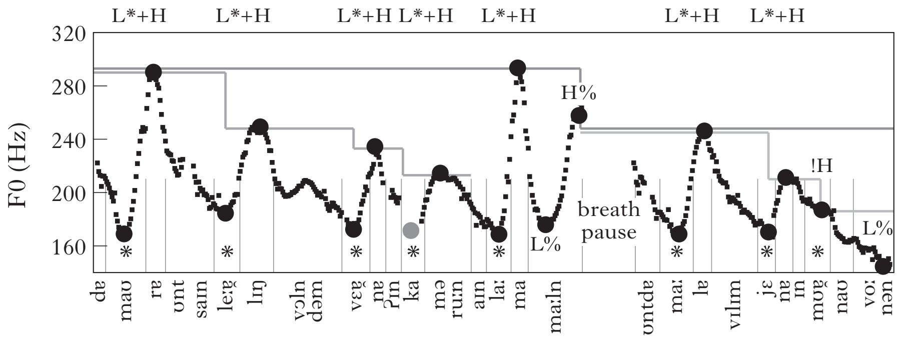


Figure 3

Annotated F0 track for [*Der MAUrer und sein LEHRling wollen dem WERner in KAmerun ein LAma malen*] [*und der MAler will im JÄNner 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.' Speaker SW. Labels are by edges of lexical words and of accented syllables. Accented syllables are marked with *. The two clauses of the utterance are separated by a breath pause. After medial downstep in the first clause, the nuclear accent of the first clause shows a rise to approximately the initial height. The reference lines indicate the phonetic analysis in terms of embedded register levels. The exact location of the L tone indicated by the grey circle cannot be directly inferred from the F0 track.

has the utterance as its domain. Downstep, on the other hand, can be undone utterance-internally, with a return to higher values. Pierrehumbert & Beckman (1988: 60, 70ff) analyse the renewed height initially in intermediate phrases in Tokyo Japanese as undoing preceding downstep, but not undoing declination. Since the lowering in Fig. 2 may be undone later in the utterance, as will be seen, it is plausibly classified as downstep, not declination, under these theoretical assumptions. Another reason for not attributing the lowering in Fig. 2 to declination is quantitative: declination is usually smaller in its effect, while the dramatic amount of lowering in Fig. 2, typical for this speaker and others, is more plausibly attributed to downstep.

An alternative source of phonetic lowering is FINAL LOWERING. In the English data of Liberman & Pierrehumbert (1984), and in the Mexican Spanish data of Prieto *et al.* (1996), downstep and final lowering interact in the following way: successive downstep leads to successive lowering of H tones. The last of these in a sequence is subject to final lowering, which leads to an unexpectedly low value in final position, relative to the course of downstep seen in the preceding values. If German had final lowering but not downstep, the lowering in the values of Fig. 2 might possibly be explained. However, longer sequences of pitch accents show successive downstep. Successive downstep cannot plausibly be attributed to final lowering. This can be seen in the Northern German findings of Grabe

(1998), and is likewise true of my Southern German data, as will be apparent in some parts of more complex contours analysed below.

In summary, downstep among H tones is analysed in terms of lowering of an abstract phonetic register. Downstep applies in German – the lowering among H tones that is observed is not plausibly attributed to declination or final lowering, with no contribution of downstep.

After this introduction to some elements of intonational analysis, I now turn to more complex cases and to the contribution of this article.

2 German upstep and its analysis

The phenomenon at the core of this article is illustrated with the F0 track in Fig. 3. By way of showing a preview of the analysis, Fig. 3 also includes schematic register levels that will be crucial in the analysis.[7] The utterance in Fig. 3 consists of two i-phrases, separated by a breath pause.[8] The rises in connection with the stressed syllables are analysed as L*+H throughout except in utterance-final position. Successive downstep is observed on the first four peaks in Fig. 3. This is followed by a return to the initial height on the fifth peak, the one on the nuclear stress of the first i-phrase. I here call this phenomenon UPSTEP. The nature of this upstep is the focus of the present investigation. Figure 3 illustrates the main properties of the phenomenon of upstep that will be experimentally confirmed in §3 below. For one thing, upstep is a return to the initial height. This also means that it disregards preceding downstep. For another, the upstepped peak is in turn followed by downstep.

As shown in Fig. 3, the present analysis of this phenomenon will involve two separate register lines. In the first i-phrase in Fig. 3, one register line models the downstep on the first four peaks. The other register line remains of constant height during the first five peaks, and thus provides the phonetic height to which the upstepped value returns on the fifth peak. However, even the register line that is constant during the first i-phrase is then downstepped between the first and second i-phrases, modelling the step of downstep that follows the upstepped peak.

In the following, I lay out the phonetic analysis of the German upstep phenomenon central to this article in the context of the literature on partial reset, and discuss a similar phenomenon in English and German, mentioned in the literature.

Standardly, RESET is taken to be the phenomenon in which previous downstep is interrupted, and a return to a higher height is observed. However, the phenomenon of upstep in Fig. 3 differs in two crucial ways

[7] [n] in [jɛnɐ] is ambisyllabic in standard accounts of German syllable structure, but is labelled as part of the following syllable.

[8] Of the four speakers evaluated below, three (MG, SW and TL) showed long breath pauses in this position with considerable regularity; the fourth speaker (CB) also frequently showed breath pauses in this position. For all speakers, internal breath pauses in any other position were rare and exceptional.

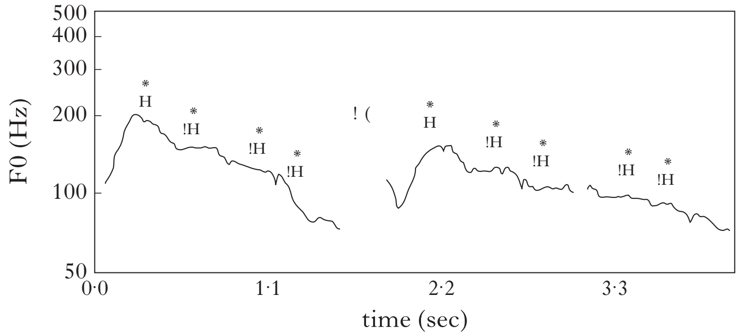


Figure 4

Partial reset in the Dutch utterance (*Merel, Nora, Leo, Remy*), *en* (*Nelie, Mary, Leendert, Mona en Lorna*). The reset at the beginning of the second phrase is discernible in the renewed height relative to the last accent at the end of the first phrase. The reset is partial rather than complete in not reaching the height of the utterance-initial peak. From van den Berg *et al*. (1992: 344).

from the phenomenon of reset. For one thing, reset is normally observed at the *beginning* of a new domain, while upstep in Fig. 3 occurs at the *end* of a domain, before a new domain begins. Thus, if the renewed height in Fig. 3 were the phenomenon of reset, it would be on the wrong side of the breath pause. Furthermore, reset – though it may be prototypically thought of as a return to the initial register – does not in fact seem to usually go up as high as that. Instead, utterance-internal reset seems to typically be PARTIAL RESET. Partial reset is illustrated for Dutch with a plot from van den Berg *et al*. (1992) in Fig. 4. It can be seen that the beginning of the second phrase here is raised relative to the downstepped values at the end of the first phrase (thus some sort of reset has applied initially in the second phrase). However, the phonetic value observed initially in the second phrase is lower than that seen initially in the first phrase (hence *partial* reset).

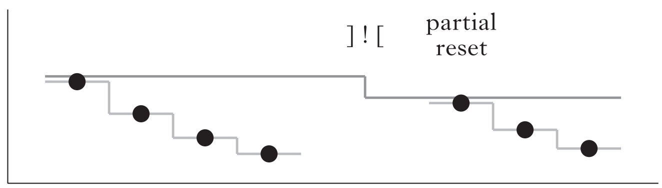
Systematic instances of partial reset are found in the study of English intonation by Ladd (1988), in the study of Yoruba intonation by Laniran (1992) and in the study on Dutch intonation by van den Berg *et al*. (1992), from which Fig. 4 is taken.[9] The second difference between the phenomenon of upstep in Fig. 3 and the phenomenon of reset as known from other languages thus seems to be that the upstep phenomenon in Fig. 3 is a return to the initial height, while reset seems to typically return only partway to the initial height across languages.

The models that have been developed for understanding partial reset

[9] In Japanese, the situation seems to be more complex. The data of Selkirk & Tateishi (1991) suggest that some speakers reset only partially, while others would seem to use a complete reset at least in some cases.

are, however, important for understanding upstep in Fig. 3. Relevant here is an idea of Ladd (1988, 1990, 1993b) for understanding partial reset, and the implementation of this idea in the model of van den Berg *et al.* (1992). Ladd proposes that a smaller, local downstep may be embedded within larger downstep. Applied to partial reset as in Fig. 4, this view holds that the second domain is downstepped as a whole relative to the first domain. Further downstep applies embedded within each domain. The resetting is partial, rather than complete, because it targets a register level which is, as a whole, downstepped relative to the first domain as a whole.

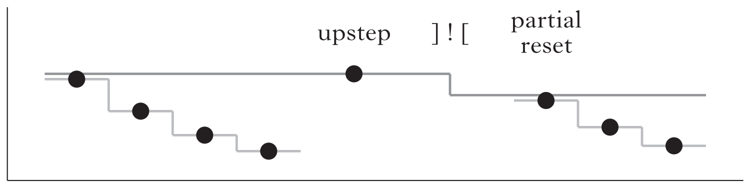
Van den Berg *et al.* (1992) strengthen this account by arguing against a boosting analysis of renewed height in the reset, relative to the preceding peak. They also show that in Dutch, downstep at the larger level is quantitatively smaller than downstep among adjacent accents, a finding to which I will return. Furthermore, they suggest an implementation of this idea that will be crucial in the analysis of the upstep here. The implementation postulates a phonetic 'phrasal reference line', here called PHRASAL REGISTER, for each of the larger domains, with the second phrase downstepped relative to the first. This is shown by the upper grey line in (3).

(3)


Embedded downstep applies within each phrase among the accents. Initially in each phrase, the embedded register is set back to the register of the phrase. Given the downstep relation among the two phrasal registers, this then amounts to partial reset at the beginning of the second phrase, as shown. I should note that van den Berg *et al.* (1992) do not employ the notion of register for the embedded downstep among the accents. Instead, they formalise a reference line only for the relation between the two large phrases, and implement downstep among accents directly, without recourse to a separate reference line. I will use a minimal modification of this model – retaining the original suggestion of Clements (1979), I will generally model downstep by changes in register. (This will be important in the suggestion on the phonology–phonetics relation in §4.) As shown in (3), the minimal modification thus assumes two separate register lines. These are here assigned to prosodic levels: the register line that is lowered once per i-phrase is taken to be the register of the i-phrase, and the register line that is lowered once per p-phrase is taken to be the register of the p-phrase.

The motivation of van den Berg *et al.* (1992) for introducing the phrasal register is to have a means of construing all downstep relations as local in some sense. Thus, on the surface, partial reset is 'downstep at a distance', between the initial peak in the first phrase and the initial peak in the second phrase. The abstract construct of the phrasal register allows an analysis in which this relation of lowering is computed locally under adjacency as a lowering relation between the register of the first phrase and the register of the second phrase.

German upstep provides independent evidence for this abstraction – upstep is claimed to be a phenomenon by which a tone late in the i-phrase returns to the phrasal register before the phrasal register is downstepped. Upstep thus may be seen as a witness to the register line of the phrasal register. The phrasal register must still be at its original height at the end of the phrase, where upstep targets this register line. This analysis is schematically illustrated in (4).

(4)


In Fig. 5, I repeat Fig. 3, annotating upstep and partial reset in this analysis. The analysis is supported by a number of properties of upstep, which are experimentally established below. Upstep shares with reset that it is a return to a register level that is independent in its height from the preceding register level. In the domain-initial partial reset, this is a return to the (downstepped) phrasal register. German upstep is likewise a return to the phrasal register, and thus likewise independent in its height from preceding downstep. Further, the two differences between upstep and domain-initial partial reset are now related to each other: since the upstep phenomenon occurs before the juncture of the i-phrases, it returns to the phrasal register before the phrasal register is downstepped. Upstep therefore represents a *complete* return to the initial height. Partial reset, occurring after the juncture, targets the already downstepped phrasal register, and is thus only a *partial* return towards the initial height.

As the annotation in Fig. 5 highlights, this German contour also contains a peak that exemplifies partial reset: the initial peak in the second clause. To appreciate this point, the reader is invited to ignore, for a moment, the upstepped peak in Fig. 5, and to compare the relation of the initial peak in the second i-phrase with the first four peaks in the first i-phrase in Fig. 5. It can be seen that the initial peak in the second clause is higher than the third and fourth peaks in the first clause – this shows how the initial peak of the second clause is reset relative to the downstep occurring earlier in the first clause. At the same time, the first peak in the second clause is clearly partial reset, not complete reset, in that its height is one register step below the utterance-initial peak (and the upstepped peak).

Let us then put together the analysis of upstep in Fig. 5 with that of partial reset initial in the second i-phrase. The upstepped nuclear pitch

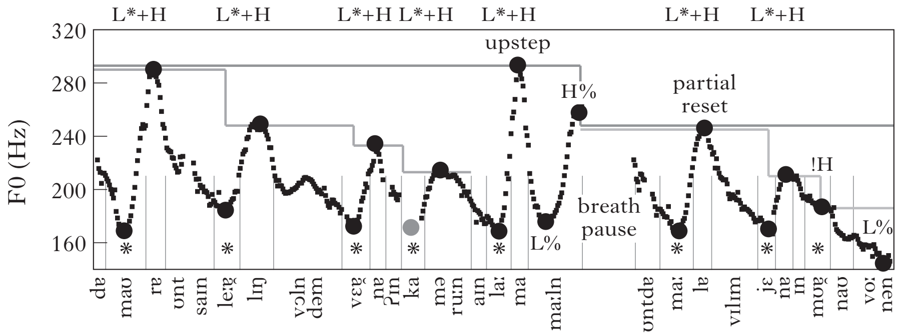


Figure 5

Repetition of Fig. 3, with upstep and partial reset indicated. In the analysis, the upstepped peak and the peak in partial reset are both scaled to the register of the i-phrase, and are separated by downstep of this register.

accent of the first clause is scaled to the phrasal register just before it lowers. The first pitch accent in the second clause is on the phrasal register just after it lowers. Thus, the downstep relation between the upstepped nuclear pitch accent and the peak in partial reset in the second clause shows us, overtly and on adjacent pitch accents, the downstep relation of the phrasal register postulated in the model of van den Berg *et al.* (1992). In other words, downstep at a distance (partial reset) in Dutch, English and Yoruba is here claimed to become downstep under adjacency in the presence of upstep in this variety of German.

A phenomenon comparable to the upstep observed here has been noted for English in single-clause utterances. According to Ladd (1983a: 735), downstep in English may either affect prenuclear and nuclear tones, or be limited to the prenuclear ones. Ladd illustrates this with the schematic contours reproduced here as (5).

(5) a.

I really don't think he ought to be doing that
H !H !H !HL L%

b.

I really don't think he ought to be doing that
H !H !H HL L%

Ladd (1983a), who develops a more general theory of downstep by featural representation, suggests that the downstep feature repeats over the entire domain in (a), and over a domain preceding the nuclear accent in (b), in a manner resembling the behaviour of vowel quality features in vowel harmony.

Beckman & Pierrehumbert (1986: 298f) discuss a comparable contour. Their example has downstep among the first three accents, which does not carry over to the final accent in narrow focus. They suggest that the final focus might introduce an intermediate phrase boundary to its left, which could then block downstep from affecting the focused element.

Insofar as it relates to focus, this phenomenon has also been observed in German. Thus, Féry (1993: 159f) notices for sentences with two H*+L pitch accents that they may be downstepped, in which case the sentence as a whole is understood as focused. Where the second accent is not downstepped, the material marked by the first accent is understood as given and is thus excluded from the focus. This case seems to share with the English example of Beckman & Pierrehumbert (1986) that narrow focus on the nuclear pitch accent blocks downstep. Narrow focus that blocks downstep was also observed on sequences of two L*+H pitch accents in Berne Swiss German by Fitzpatrick-Cole (1999). Across English and German, it would seem that continuation of downstep on a final nucleus is in some sense the default, and that narrow focus on the nucleus can block this default. (5b) suggests that English also allows the discontinuation of downstep on the nucleus in the absence of narrow focus on the nucleus. The choice of example suggests that this might occur in emphatic or expressive contexts.

It seems likely that the phenomena reviewed in this section are closely related to upstep in German, and may likewise be returns to the phrasal reference line on the nuclear pitch accent. If so, the analysis would not require the assumption of a separate domain preceding the renewed height in the suggestions outlined by Ladd and Beckman & Pierrehumbert (1986).[10]

The analysis in which a tone late in the clause returns to a high phonetic reference line after clause-internal lowering has an antecedent in Bruce & Gårding's (1978) suggestions on the scaling of the high 'sentence accent' in some Swedish dialects. See Truckenbrodt (2001) for more discussion of this parallel.

3 Experimental findings

This section presents the experimental results from four German speakers who show upstep on the nuclear pitch accent. The discussion seeks to establish the correctness of the register analysis in §2. It is argued that upstep is a return to the initial height, that it neutralises preceding downstep and that it stands in a downstep relation to a following partial

[10] A problem with this additional domain might be that it does not appear to be a stress domain, i.e. there does not appear to be a strongest stress defined relative to it.

reset. §3.1 describes the stimuli, notation, recordings and criteria for measurements in the experiment. §3.2 discusses the results.

3.1 Description of the experiment

3.1.1 *The stimuli.* The stimuli were constructed so as to reliably elicit certain prosodic patterns from naive subjects with a minimum of instruction. This was accomplished through control of focus, as well as choice of syntactic structures that would lead to regular and predictable assignment of prosodic structure.

The experiment was designed to elicit the two-clause stimuli under maximally wide focus (or, if you will, no focus at all) (Rooth 1992, Schwarzschild 1999). This was done by eliciting each sentence as an answer to the question *Was gibt's Neues?* 'What's new?', read by the experimenter before the sentence was then read by the subject. This set-up was reinforced by the instructions described in §3.1.2 below.

Within the wide focus thus elicited, the stimuli consisted of two coordinated CP clauses, as shown in (6) in the last line. The two CP clauses in the two-sentence condition constituted two i-phrases, as is shown in the first line of (6). This follows the cross-linguistic pattern observed in Nespor & Vogel (1986) – that root clauses are separate intonation phrases. At the end of the first i-phrase the stimuli contained an unscrambled direct object followed by a main verb in clause-final position ([… DO V]$_{CP}$ in (6)).[11]

(6) [x]$_{I1}$ [x]$_{I2}$
(x)(x)(x)(x)(x)$_{P}$ (x)(x) (x)$_{P}$
L*+H L*+H L*+H L*+H L*+H (…) H% L*+H L*+H (…) L* L%
[SUBJ (& SUBJ (IO (PP))) DO V]$_{CP}$ [XP XP XP]$_{CP}$

The descriptions and accounts of German stress agree that nuclear stress (and thus the final pitch accent) of a clause or i-phrase falls on the unscrambled direct object in such cases, so long as effects of narrow focus do not interfere (Schmerling 1976, Höhle 1982, Gussenhoven 1983, 1992, Krifka 1984, 1998, Pheby 1984, Selkirk 1984, von Stechow & Uhmann 1986, Uhmann 1991, Cinque 1993, Féry 1993, Jacobs 1993, Büring 2001, among others). The last line of (6) also shows the major syntactic constituents that preceded the nuclear stress in the first clause. They formed p-phrases, and carried phrasal stress (second line) on which pitch

[11] The control of focus in the stimuli was also designed to prevent scrambling of the direct object or other arguments. For example, a definite direct object will scramble iff it is contextually given (Jäger 1995), a condition that extends to proper names, while indefinites scramble if they are presuppositional (Diesing 1992). The all-new context discouraged the assumption that the direct object would be contextually given or presuppositional, and so lead to a rendition with the direct object unscrambled. This is important for stress assignment, in that a scrambled direct object will lead to stress assignment on the following verb (Cinque 1993), as well as often being de-accented itself, due to its givenness.

accents were assigned (third line). The syntactic constituents are SUBJ (subject), & SUBJ (subject coordination), IO (indirect object), PP (PP adjunct), DO (direct object or VP complement) and V (verb). The placement of accent on the prenuclear constituents may be seen to follow the suggestion of Gussenhoven (1983, 1992) for English and Dutch and Selkirk (1995) for English, or the application of proposals in Truckenbrodt (1995) to German in Büring (2001): one accent is assigned per major constituent preceding the verb. The coordinated subject leads to two separate accent domains, a stable pattern in my data, as correctly predicted by an application of the accounts of Selkirk (1995) or Truckenbrodt (1995) and Büring (2001).[12] The second clauses had different syntactic realisations, and were consistently designed so as to carry three beats of stress with three accents on three XPs. The stimuli were constructed to obtain stable patterns of prosodic structure according to the schema in (6), with two i-phrases with their nuclear accents in a predictable position, and with the number of prenuclear accents flexible in a controlled way. This set-up was very successful in that the speakers produced, in a natural way and with considerable regularity, sentences with the expected prosodic characteristics.

The tonal patterns found in the recordings are phonologically analysed as shown in the third line of (6).[13] Notice that all pitch accents except for the utterance-final one are L*+H. The utterance-final pitch accent was typically H+L* for SW and TL, and was measured as L* for CB and MG, with only occasional evidence for H+L* in the recordings reported on here. All speakers showed evidence of an utterance-final L% edge tone, and of a H% edge tone at the right edge of the first (i.e. non-final) i-phrase. Some speakers showed evidence for an additional edge tone preceding H% that will be discussed shortly.

(6) subsumes four sets of stimuli with different length. In the longest set, all syntactic major constituents shown in the last line of (6) were present. Three shorter sets did not contain some of the syntactic constituents, as shown by the parentheses in the last line of (6). The four sets thus had different numbers of pitch accents in the first clause, as shown in (7). (7) also introduces notation that will be used in the rest of this article. H_1 and H_2 are the first two H tones of the utterance. Single underlining is used to highlight the L*+H pitch accent in immediately prenuclear position in the first i-phrase, or its H part. Double underlining is used to highlight the nuclear L*+H pitch accent in the first i-phrase,

[12] Uhmann (1991) assumes that such prenuclear accent domains are generally optional in German. However, I have found, both in pilot recordings and in the present study, that such prenuclear accents are regularly assigned when occurring within wide focus, as predicted by the theories of Gussenhoven (1983, 1992) and Selkirk (1995).

[13] Here H_P edge tones are omitted. Most of the stimuli left two syllables for the execution of L*+H H_P, so that +H and H_P usually temporally coincide and are indistinguishable. Since they are empirically of comparable height when both occur in proximity to each other (see also Truckenbrodt 2001), the distinction between +H and H_P is not relevant here.

or its H part. Further, H′ and H″ are the first two H tones in the second i-phrase.

(7) a. x x
[x x]$_{I1}$ [x x x]$_{I2}$
L*+H$_1$ L*+H$_2$ (L%) H% L*+H′ L*+H″ (H+)L* L%

b. x x
[x x x]$_{I1}$ [x x x]$_{I2}$
L*+H$_1$ L*+H$_2$ L*+H (L%) H% L*+H′ L*+H″ (H+)L* L%

c. x x
[x x x x]$_{I1}$ [x x x]$_{I2}$
L*+H$_1$ L*+H$_2$ L*+H L*+H (L%) H% L*+H′ L*+H″ (H+)L* L%

d. x x
[x x x x x]$_{I1}$ [x x x]$_{I2}$
L*+H$_1$ L*+H$_2$ L*+H L*+H L*+H (L%) H% L*+H′ L*+H″ (H+)L* L%

Example stimuli are shown in (8). In their rendering here, they are shown with capitalisation to indicate phrasal stress, and with underlining and double underlining to highlight, respectively, the immediately prenuclear and the nuclear stresses in the first clause. Further, the clauses are separated by a line break. These devices were not employed in the presentation of the stimuli in the experiment.

(8) 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. Die LEna und die MAnu wollen die NONne malen,
und der WERner soll in MURnau einen RoMAN schreiben.
'Lena and Manu want to paint the nun,
and Werner is supposed to write a novel in Murnau.'

c. Der WERner und die LEna wollen der NONne ein LAma malen,
und der HEIner will in HAMburg eine MOle mauern.
'Werner and Lena want to paint a llama for the nun,
and Heiner wants to build a mole in Hamburg.'

d. Die MAnu und die HANne sollen der LEna im JAnuar das LEInen weben,
und der WERner soll in MURnau MaROnen holen.
'Manu and Hanne are supposed to weave the linen for Lena in January,
and Werner is supposed to get sweet chestnuts in Murnau.'

The complete set of stimuli consisted of six different sentences for each of the four length patterns in (7). These were presented in a pseudo-randomised list along with a comparable number of shorter sentences. The list was read three times by each subject, so that, for each subject, a total of 18 tokens of each of the four patterns in (7) could be evaluated.

As far as possible, the stimuli were constructed as all-sonorant utterances, with particular attention to the positions of anticipated H tones. This was to minimise consonantal interference with the F0 values. High front vowels were also avoided in the expected positions of H tones. Likewise avoided were 'vowel-initial' words, due to the effect of a glottal stop in this position on F0. The nouns chosen were usually bisyllabic with initial stress (in some cases trisyllabic with initial or penultimate stress, or monosyllabic with a sonorant coda). That was so as to allow for a full execution of the H in L*+H.

3.1.2 *Recordings*. The recordings with SW reported here were made in Cambridge, Mass., following up on earlier recordings I had made with SW. The other speakers discussed here are from a set of speakers who were recruited at the University of Tübingen, and were reimbursed for their time. Of these speakers, four were later discarded: one because there was no regularity in the scaling of the L*+H pitch accents (in particular, no downstep or upstep), and two others because they preferred to read the two clauses of the stimuli as separate, without upstep intonation that announces the second clause at the end of the first. A fourth speaker was discarded because she grew up bilingual. The remaining speakers grew up with German as their only native language. The data of three of these speakers are included in the following report, along with the data of SW. The data of four other speakers, who did not show upstep on the nuclear pitch accent, are evaluated in Truckenbrodt (2001) in connection with an upstep phenomena on their medial H% edge tones. That investigation also includes an evaluation of the scaling of the H% edge tones of the four speakers discussed here.

Speakers received instructions in connection with emphasis (focus), instructions in connection with choice of melody, and procedural instructions. The instructions were as short as their rendition here. The instructions related to focus were intended to reinforce the elicitation of sentences with wide focus (see the preceding section). It was explained that foregrounding by particular emphasis on any particular part of the sentence was not of interest, that all parts of the sentence should be taken to be equally important, and that the sentences were to be read as answers to the question *Was gibt's Neues?* 'What's new?' (read by the experimenter before each token) so as to encourage that assumption. It was also said that each stimulus should be read as though standing on its own, such that if a particular character appeared in more than one stimulus, that character should be assumed to be new each time. This set-up proved to be highly effective for all subjects. The subjects were able to follow these instructions with what appeared to be natural ease.

Further instructions to the speakers were used to regularise the choice of tonal melody. The instructions did not include an illustration of how the sentences were to be read. However, it was briefly explained how sentence melody can be used to express various things such as incredulity, and an illustration of a short sentence with a final rise was pronounced. It was explained that such expressive intonation was not of interest, and that, instead, a normal narrative reading of the sentences was desired.

In the procedural instructions, speakers were told to ask for a repetition of a sentence, including question and answer, if they felt that their reading had not been natural, or if their reading of a sentence wasn't fluent. All speakers made some use of this option. In those cases, only the last of a series of repetitions was analysed.

The sentences were presented as computer print-outs in the form of a pseudo-randomised list that contained each sentence once. The text included, for each sentence, the question *Was gibt's Neues?*, followed by the sentence to be read as an answer. The question was read by the experimenter, and the answer sentence by the subject. Initially, about ten question–answer sequences were read without recording. This was helpful for making the subjects comfortable with the process. Subsequently, the recordings were made. The list was read three times, with short breaks between readings. The actual recordings lasted between 20 and 30 minutes for each subject.

All recordings were made on a DAT tape-recorder in a quiet room. They were later recorded into a Sun computer for analysis with Entropic speech-processing software (ESPS/*waves+*). For each recording, the analysis included labelling of word boundaries with the help of spectrograms, checked against an auditory impression, and measurements of the F0 track by the criteria described in the following section.

3.1.3 *Criteria for F0 measurements*

3.1.3.1 *Exclusion of values.* A measurement at a given point was skipped iff one of the following four circumstances was found. (i) A speaker exceptionally used a different pitch accent from the one he or she regularly used in that position. Though this sometimes occurred on the utterance-final pitch accent for other speakers in my study, the speakers reported on here did this extremely rarely. (ii) F0 distortion due to a consonant in the position of the expected F0 maximum or minimum was so considerable as not to allow a measurement. This is usually limited to the positions of L tones. (iii) There was no evidence in the F0 curve of the presence of a tone, i.e. when the local F0 curve could be accounted for by linear interpolation between assumed preceding and following tonal values. For example, the expected second H for speakers who used the final contour L*+H H+L* was sometimes not discernible, in which case no measurement of it was made. (iv) The F0 routine of the software did not find any plausible pitch points in the relevant location, which was sometimes the case with L% in utterance-final position. Table I shows how many measurements are missing due to these criteria. They are broken up into measurements of the utterance-final sequence (H)+L* L% and all others (all non-utterance-final L*+H pitch accents and H%

or L % H % sequences at the medial boundary). Crucial is the first row in Table I that shows the missing values of non-utterance-final tones, the ones chiefly considered in this investigation. Between 0 % and 0·6 % of the measurements are missing in positions other than utterance-final. More measurements are missing in utterance-final position, due to questions of discernability of the second H in the final sequence L*+H H+L* and due to the occasional problems with determining the height of the utterance-final L %.

		SW	CB	TL	MG
non-U-final	%	0·4	0·3	0·6	0·0
	abs.	4/900	3/864	5/864	0/864
U-final	%	14·4	17·4	15·7	2·8
	abs.	31/216	25/144	34/216	4/144

Table I

Missing values, with utterance-final values ((H)+L* L%) separated. Few values are missing in earlier positions, on which the present investigation concentrates.

The totals vary because of the different numbers of edge tones preceding the medial edge of the i-phrase (L % H % *vs*. H %), and different choices of pitch accents in utterance-final position.

Further, there were no missing values at all for any of the four speakers in the position of upstep on the nuclear pitch accent, the focus of the present study.

3.1.3.2 *The point of measurement*. A theory of the phonology–phonetics relation is assumed in which a tone that is phonologically associated with one syllable is not necessarily found in that syllable phonetically (Silverman & Pierrehumbert 1990, Myers 1999, Ladd *et al*. 2000 and references there).Typically, the L portion of a L*+H pitch accent leads to an F0 minimum around or shortly after C-V boundary of the stressed syllable in my data. However, the measurements located L* as the local minimum in that area, which included the possibility of L* being further inside of the stressed syllable or further to the left, including positions minimally preceding the stressed syllable. As in English (Pierrehumbert 1980), the H portion of L*+H does not appear to be limited to occurring in a particular syllable. In the German recordings, if the stressed syllable is long, +H typically occurs within the stressed syllable, and if the stressed syllable is short, +H will typically occur in the syllable following the stressed syllable. +H in prenuclear position was therefore consistently measured as the local maximum in the area between two L* tones. In the typical case, this was found at the end of the rise from the preceding L*. A more complex situation arose in the case of the plateau for speakers CB, MG and TL in nuclear position in the first clause,

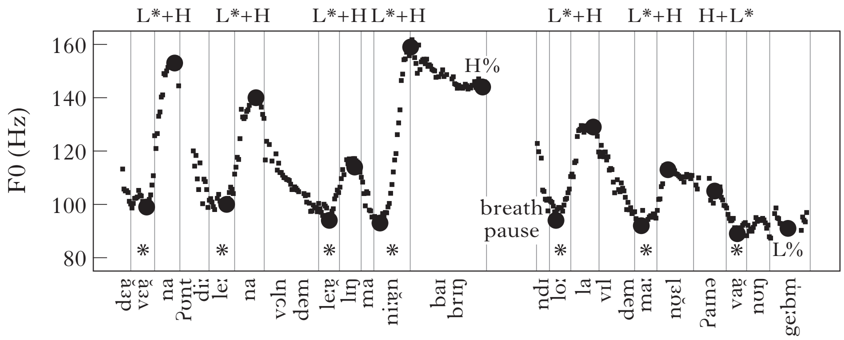


Figure 6

F0 track for [*Der WERner und die LEna wollen dem LEHRling MaNIEREN beibringen*] [*und die LOla will dem MAnuel eine WARnung geben*.] 'Werner and Lena want to teach manners to the apprentice, and Lola wants to give a warning to Manuel.' Speaker TL. The upstep on the nuclear rise of the first clause is followed by a high plateau up to the breath pause at the end of the first clause.

analysed as spanning the interval between the upstepped L*+H and a H% at the edge of the i-phrase. In the presence of a small dip in the plateau, +H was measured as the highest point preceding the dip, and H% as the highest point following the dip, which then regularly occurred close to the edge. Without a dip, +H was measured in the highest point towards the left edge of the plateau, i.e. after the rise had fully taken effect, and H% was measured in the highest point towards the right edge of the plateau, before any fall at the end of the plateau set in (such a fall sometimes occurred, beginning in the final syllable of the i-phrase).

In some cases, there were obvious consonantal distortions of the F0 curve, though local enough to allow for plausible F0 measurements outside of the distortions. This arose almost only with L tones, as the stimuli had sonorous segments in the expected places of H tones almost throughout. In these cases, the F0 points due to the consonantal distortions were ignored, and the measurements were taken in a part of the F0 curve not thus affected. Likewise ignored were F0 points that were outliers, i.e. points that lay outside of the otherwise regular and contiguous F0 curve.

3.2 Results

3.2.1 *Intonational patterns*. The illustrations above are from speaker SW. This speaker grew up in Villach, Austria, and later lived in Vienna before moving to the U.S. for graduate studies. I made numerous recordings with SW before the recordings reported here, and the medial nuclear pattern was constant across recordings with two i-phrases. The recordings with SW reported below are from a single session. As can be seen in (4) above, the upstepped nuclear pitch accent of SW is followed by a combination of L%H% edge tones at the right edge of the i-phrase. The three other speakers discussed here, CB, TL and MG, had a slightly different tonal pattern following the upstepped nucleus, which is illustrated in Fig. 6. This pattern features an upstepped plateau that is plausibly analysed as spanning the interval between the upstepped nuclear L*+H and a following H% edge tone, likewise on the upstepped level.

The three speakers of this medial pattern in my sample, CB, TL and MG, are from different parts of Baden-Württemberg. However, no dialectal classification is offered here. Other medial patterns are also used in Baden-Württemberg (Truckenbrodt 2001).[14] Further, Grabe (1998: 82ff) observes individual variation following L*+H in nuclear utterance-medial position with speakers of Northern Standard Germany (Westphalia), of which one seems to be the rise followed by a plateau seen here, and another may be comparable to the one reported for the other Southern speakers in my sample in Truckenbrodt (2001).

Figure 7 shows plots of the measurements of the four speakers. For each speaker, the four sets are plotted separately, such that each line shows the averaged measurements from 18 recordings of a set (with missing values as described in §3.1.3.1). The four sets are plotted left to right up to the upstepped rise, and are plotted together for the remaining tones of the utterance. The four sets in Fig. 7 are labelled (a)–(d). Notice that upstep, highlighted by circles, immediately follows the initial peak in set (a), and is preceded by one, two and three steps of downstep in sets (b), (c) and (d), respectively. The plots give a sense of the consistency of some of the main characteristics of the intonational patterns. First, downstep among prenuclear accents is regularly found. Second, upstep on the nuclear accent of the first clause regularly occurs for all speakers in all four sets. Third, the plots give an impression of the height of the upstepped tones, which is generally comparable to the height of the utterance-initial peaks. Fourth, the plots give a first assessment of the neutralising effect of upstep. TL, for example, shows upstep of almost identical height in all sets, exhibiting complete neutralisation of preceding downstep in upstep. MG, on the other hand, shows a reflex of preceding downstep in the upstepped peaks: sets with more downstep preceding upstep show a somewhat lower upstep value than sets with less downstep preceding upstep. Here the neutralisation is incomplete. For this and arguably related reasons, the data of MG will be discussed separately below. Fifth, upstep of SW in each set is followed by the values of the L%H% edge-tone combination, while speakers TL, CB and MG show a high plateau under upstep. Finally, across speakers, the first peak in the second i-phrase, on H′, is generally lower than the upstepped

[14] TL lived in Waldshut, in Lörrach and in Sasbach near Achern before he was 18, and CB grew up in Zogenweiler near Ravensburg. MG comes from a suburb of Reutlingen.

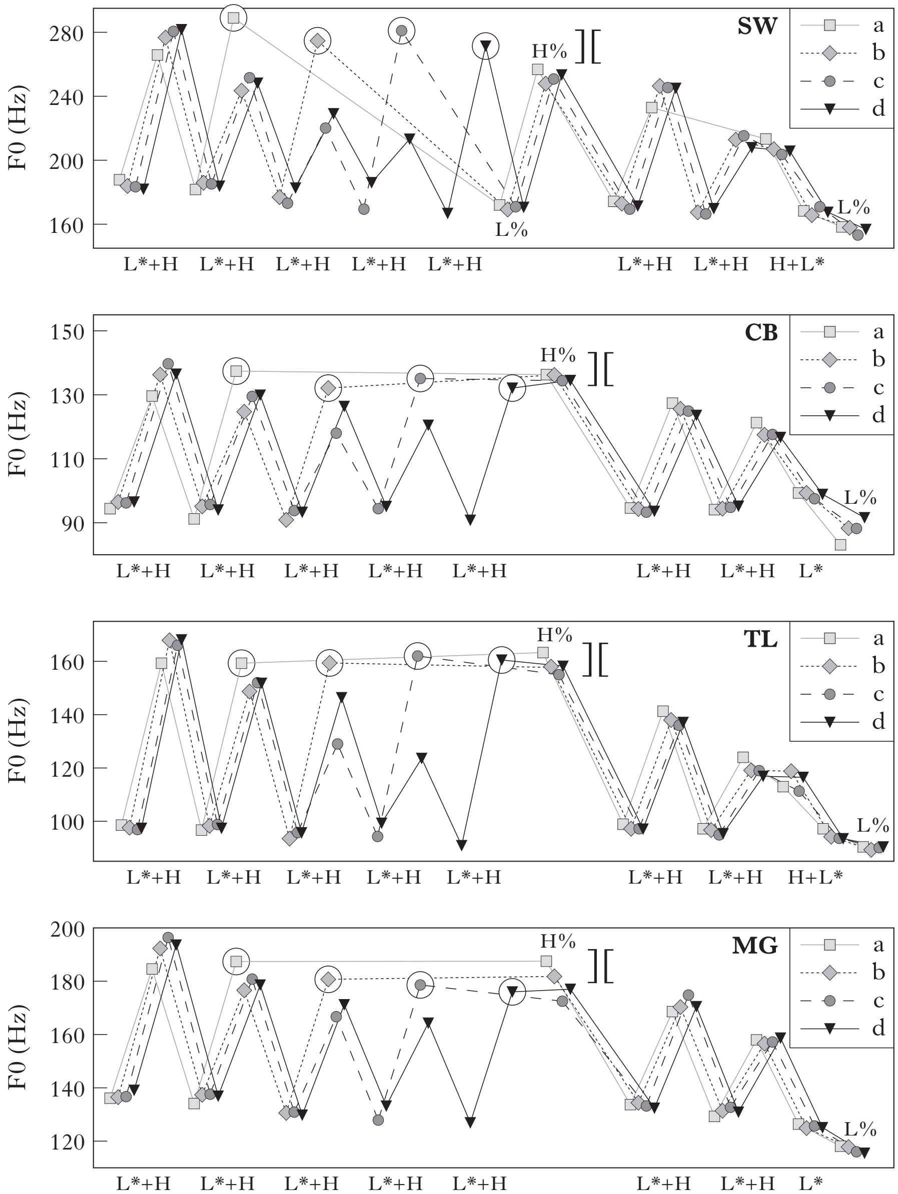


Figure 7

Average values of L and H points in four sets of different length (a)–(d). Each point is an average of nominally 18 measurements. The points of upstep on the nuclear rise in the first clause are circled.

value. Recall that the present analysis treats this as downstep of the phrasal register.

In the following sections, the measurements are evaluated in more detail.

3.2.2 *Height as initial*. One source of support for the analysis of upstep as a return to the phrasal register comes from a comparison of the upstepped values with the utterance-initial peaks. Table II shows such a comparison in some detail. As the results of a t-test shows, the difference between H_1 and upstepped $\hat{}\underline{\underline{H}}$ is not significant for either SW or CB. For speaker TL, the upstepped $\hat{}\underline{\underline{H}}$ is somewhat lower than H_1, in a way that is statistically significant. The shaded column shows the size of the actual average upstepped rise ($\hat{}\underline{\underline{H}} - \underline{\underline{L}}*$) in relation to what would be a rise to the exact same height as the initial peak ($H_1 - \underline{\underline{L}}*$). The difference is only 3 % for SW and CB. Even though TL does not have a full return to the initial height, this calculation shows that the upstepped rise of TL nevertheless measures 93 % of the expected, fully upstepped, rise.

	U-initial H_1 (Hz)	upstepped $\hat{}\underline{\underline{H}}$ (Hz)	t-test	($\hat{}\underline{\underline{H}}-\underline{\underline{L}}*$)/($H_1-\underline{\underline{L}}*$)	*inter-pretation*
SW	276 (12·4)	279 (18·5)	ns	1·03	upstep, no declination
CB	136 (5·9)	134 (5·8)	ns	0·97	
TL	165 (10·6)	160 (8·2)	***	0·93	upstep, declination

Table II

Comparison of average values of the utterance-initial peak (H_1) and the upstepped peak ($\hat{}\underline{\underline{H}}$). Standard deviation (in Hz) is shown between parentheses. Paired sample t-tests show no significant difference for speakers SW and CB, and significant differences for speaker TL. The penultimate column shows the upstepped rise in relation to what would be a full return to the initial height. The comparison suggests that SW and CB execute a full return to the initial height, while TL shows an overlaid effect of declination.

Taken together, these numbers show that SW and CB have a transparent return to the initial height. The somewhat lower upstepped values for TL are quantitatively within the range of what one might expect from speaker-specific declination applied to the present case (see Pierrehumbert 1980 on English, Pierrehumbert & Beckman 1988 on Japanese and §1.2 above). The fall is small in absolute terms (5 Hz) and occurs across a stretch that is on average somewhat shorter than half of the entire utterance.

In summary, the comparison between upstepped values and utterance-initial peaks supports the analysis of upstep as a return to the phrasal register. The upstepped values are plausibly understood in terms of a return to the initial register, overlaid by a speaker-specific effect of declination for TL.

3.2.3 *Neutralisation of preceding downstep*. Another source of support for the analysis of upstep as a return to the phrasal reference line comes from the neutralisation of earlier downstep distinctions in the upstepped peak. This is predicted by the present analysis, where the height of the phrasal reference line to which the upstepped values return is defined independently of the height of a preceding downstepped register. Notice that the return to the initial height and the neutralisation of preceding downstep distinctions could each in principle be false, independently of each other. It could be, for example, that the upstepped height in each set is a function of preceding downstep, but that the average in the upstep happens to resemble the initial height. That would suggest an analysis in which the upstep is a function of preceding downstep, and is not systematically scaled to the initial height. Conversely, it could be that all sets of a given speaker upstep to the same height, thus neutralising preceding downstep in the position of upstep, but that this height is very different from the utterance-initial height. That would suggest an account in which upstep is a return to an externally defined value, though not on the phrasal reference line. It is the convergence of the two predictions of the analysis, return to the initial height and neutralisation of preceding downstep distinctions, that supports the scaling of the upstepped peak on the phrasal reference line.

The neutralising effect of upstep is established with two analyses of variance of each speaker in Table III. Each analysis of variance is displayed in one row of the table. For each speaker, the first such analysis establishes the presence of significant distinctions in the height of peaks immediately preceding upstep, across sets. This shows that there are distinctions preceding the position of upstep that could in principle lead to distinctions in the following position of upstep. The second analysis of variance shows that there are no such distinctions in the position of upstep across sets.[15]

Consider first the comparison of the immediately prenuclear H̲ values (see (7)) across the sets. The most important results are highlighted in the post hoc comparisons on the right in Table III. Thus, the shaded cells indicate where downstep leads to significant distinctions in the prenuclear values across sets. For example, in the column 'a−b', it can be seen that

[15] For SW and TL, Tukey's HSD test was applied for the post hoc comparison of the analysis of variance, because the variances are homogeneous by Levene's statistic. For CG, Levene's statistic showed significant deviations from homogeneity of variance for the nuclear accent, and deviations from homogeneity of variance that approached significance for the prenuclear accent. Therefore, the post hoc statistic by Tamhane is used for both prenuclear and nuclear H for speaker CG. It gives equivalent results to Dunnet's T3 and Games-Howell in all cases for CG.

		Means (Hz)				ANOVA		Mean difference (Hz)					
		a	b	c	d	*F*	*df*	a−b	b−c	c−d	a−c	b−d	a−d
SW	$\underline{\text{H}}$	266 (15·0)	244 (13·7)	220 (11·1)	213 (8·2)	65·8 ***	3, 66	22·3 ***	23·4 ***	6·8	45·8 ***	30·3 ***	52·6 ***
	^$\underline{\underline{\text{H}}}$	289 (19·9)	275 (16·1)	280 (19·6)	271 (15·1)	3·5 *	3, 66	14·3 a	−5·7	9·5	8·6	3·8	18·1 *
CB	$\underline{\text{H}}$	130 (3·5)	125 (4·3)	118 (5·9)	121 (6·5)	17·0 ***	3, 67	4·9 **	6·7 **	−2·5	11·6 ***	4·2	9·1 ***
	^$\underline{\underline{\text{H}}}$	137 (7·4)	132 (5·6)	135 (3·5)	132 (4·5)	4·1 **	3, 67	5·6 a	−2·9	2·7	2·6	−0·3	5·3 a
TL	$\underline{\text{H}}$	159 (10·3)	148 (12·8)	129 (16·0)	123 (13·9)	27·2 ***	3, 66	10·6 a	19·5 ***	5·6	30·0 ***	25·1 ***	35·6 ***
	^$\underline{\underline{\text{H}}}$	159 (7·2)	159 (9·5)	162 (6·7)	161 (9·6)	0·38	3, 66	−0·1	−2·4	0·8	−2·5	−1·6	−1·7

Table III

Evaluation of the neutralising effect of upstep. For each speaker, the first row shows a one-way ANOVA of the immediately prenuclear values of the four sets ($\underline{\text{H}}$). Means are shown on the left. Standard deviation (in Hz) is shown between parentheses. The crucial significance results of the post hoc analysis are shown in the *Mean difference* columns on the right. By and large, the first two steps of downstep are statistically significant in this comparison across sets. For each speaker, the second row shows a one-way ANOVA of the upstepped nuclear values of the four sets (^$\underline{\underline{\text{H}}}$). By and large, the post hoc distinctions here are not statistically significant. Thus the significant distinctions in the position preceding the upstep do not lead to significant distinctions in the position of upstep.

the distinction between the prenuclear $\underline{\text{H}}$ in set (a) (first peak, not preceded by downstep) and the prenuclear $\underline{\text{H}}$ of set (b) (second peak, preceded by one downstep) is significant for speakers SW and CB (and approaches significance for TL).[16] Likewise, the shaded cells in the column 'b−c' show that the difference between the prenuclear $\underline{\text{H}}$ of set (b) (second peak, preceded by one downstep) and the prenuclear $\underline{\text{H}}$ of set (c) (third peak, preceded by two steps of downstep) is significant for all speakers. By and large, then, the first two steps of downstep are significant in this data. The difference between two *vs.* three steps of downstep ('c−d') is not significant for any speaker. This is not surprising: Liberman & Pierrehumbert (1984) argue that downstep takes the course of exponential decay. We therefore expect that its effects become smaller with the number of steps, and thus unlikely to reach statistical significance

[16] Here and below I use standard notation: a for $p < 0{\cdot}10$, * for $p < 0{\cdot}05$, ** for $p < 0{\cdot}01$ and *** for $p < 0{\cdot}001$.

as the steps become too small. The last three columns show that differences between no and two steps of downstep, between one and three steps of downstep and between no and three steps of downstep are also significant in the data, with one exception for CB.

The second part of the statistical analysis compares the nuclear ^$\underline{\underline{H}}$ values across sets. For each shaded cell representing distinctions in the prenuclear $\underline{H}$, the cell immediately below it reveals whether there are also significant distinctions in the immediately following ^$\underline{\underline{H}}$ values. For example, the white cells in column 'b – c' show that for all speakers, the distinctions between one and two steps of downstep in $\underline{H}$ are not reflected in distinctions in the nuclear ^$\underline{\underline{H}}$. Upstep has neutralised the distinctions of preceding downstep. As the table shows, most of the downstep distinctions, including the larger ones of two and three steps of downstep, are neutralised in the following upstepped values. Cell-pairs in which such significant distinctions in $\underline{H}$ are neutralised in ^$\underline{\underline{H}}$ have thick black lines around them.

The only distinction that is not fully neutralised in the position of upstep is that between sets (a) and (d) for speaker SW. Below I will motivate a process of register undershoot for speaker MG, by which the return to the phrasal register may be less than fully executed. It will be seen that this process allows a reflex of the preceding downstepped height to emerge in the upstepped values. SW appears to have a small amount of this, not large enough to attain statistical significance in all other pairwise comparisons, but still large enough to emerge in the largest distinction in ^$\underline{\underline{H}}$ values, between no and three steps of downstep ('a – d').

Overall, the statistically significant distinctions among the immediately prenuclear pitch accents are clearly neutralised in the upstepped nuclear pitch accents. This supports the analysis of upstep on the nuclear pitch accent as a return to the phrasal reference line, which predicts independence of the upstepped values from preceding downstep.

Both the preceding results speak against an analysis of the renewed height of the upstepped nucleus in terms of boosting relative to the preceding peak, i.e. amplification of the phonetic value of the preceding peak with a factor reflecting the greater prominence of the nuclear stress (Pierrehumbert 1980). Such a boosting analysis would make it arbitrary that the height of the upstepped peak is comparable to that of the utterance-initial peak. Furthermore, and more tellingly, a boosting analysis would assign the upstepped peak a value that is a function of the preceding peak. Such functions might plausibly predict that the significant differences in the preceding peaks among sets in my data should be reflected in significant differences in the upstepped peaks. If, for example, the boosted value is calculated from the preceding value by multiplication with a factor greater than one (against some reference line), then the differences in the preceding values should in fact be amplified in the boosted values. These should then turn out to be of considerably different height, with significant distinctions among sets in the upstepped values. Furthermore, if upstep were comparable to the utterance-initial peak in one of the sets, the other sets should then show considerable differences from the initial value, contrary to fact.

Kubozono (1989, 1993) argues for some effects of renewed height in Japanese in terms of a metrical boost. For Kubozono, downstep and boosting simultaneously apply to the element with renewed height. Applying this suggestion, one might wonder whether downstep, which has a partly neutralising effect, might eliminate preceding distinctions, such that boosting, even if it amplifies distinctions, is left with no distinctions to amplify. Take, for example, the distinction between sets (b) and (c). If the upstepped values first undergo abstract downstep, the distinction between these two sets will likely be statistically neutralised, resembling the case of a third, not statistically significant, downstep. When boosting applies to these abstractly downstepped values, there may be no significant distinctions to re-amplify. However, it was seen that the first two steps of downstep are statistically significant in my German data. Therefore, the distinction between sets (a) and (b), even if they abstractly undergo downstep in the upstepped peaks, should still show significant distinctions, resembling the statistically significant second step of downstep. When these values undergo boosting, their distinctions should still then be statistically significant. Thus, even an application of boosting to downstepped values could not account for the neutralisation in the upstepped peaks in my data.

3.2.4 *Upstep in relation to partial reset*. This section provides support for the analysis of upstep in the model of van den Berg *et al*. (1992), by confirming the presence of partial reset in the German materials and by establishing the relation of downstep of the phrasal reference line between the upstepped peak and the peak in partial reset (H′).

Consider the plots in Fig. 8. Only H tones are plotted. A horizontal reference line is drawn through the average values in partial reset, to facilitate comparison with the values in the first clause.[17]

Notice first that the initial value in the second clause is higher than the downstepped values in positions H_3 and H_4 in the first clause. This confirms that there is renewed height in the position of partial reset, when compared with the downstepped values in the first clause. Second, notice that the renewed height is partial rather than complete reset in being significantly lower than either the utterance-initial peak or the upstepped peak. The difference is attributed to downstep of the phrasal reference line in the present analysis.

Table IV shows a quantitative comparison of the phrase-internal

[17] The recordings with SW were made before those with the other speakers, with stimuli in which set (a) (but not the other sets) had only two pitch accents instead of three in the second clause. In set (a) of SW, the tone H′ was regularly lower than H′ in the other sets. These values are plotted with H″ rather than H′ in Fig. 8, and are also excluded in the calculations relating to H′ in the following. See the Appendix for the stimuli.

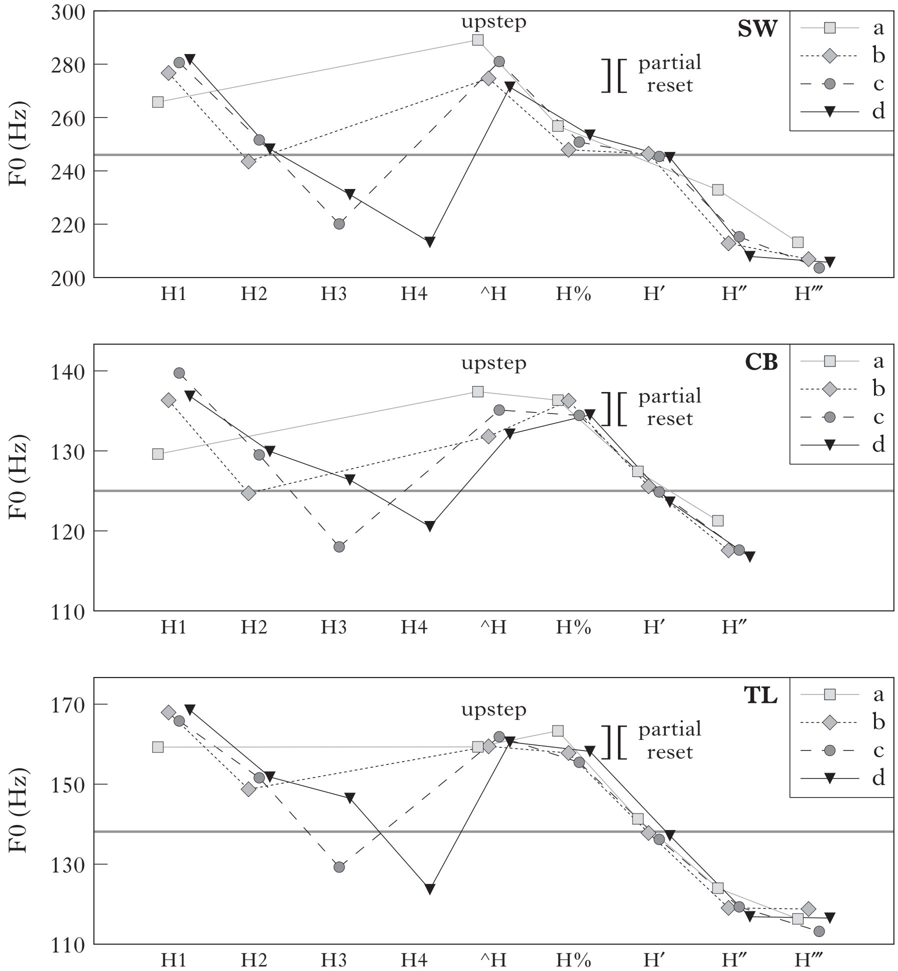


Figure 8

Average values of H points, omitting L points. A grey reference line is plotted through the average values of H′. The reference line shows that H′ is typically higher than H3 and H4, and thus reset relative to earlier downstep. It is noticeably lower than the utterance-initial peak and the preceding upstep, and thus only partially reset.

downstep relation between utterance-initial first and second peak (H_1 and H_2) with the downstep relation I claim exists between upstep and partial reset ($\underline{H}$ and H′). By the design of the experiment, there are three sets, (b)–(d) (i.e. (7b–d)), which have initial downstep between H_1 and H_2.

Table IV presents a comparison of initial downstep with medial downstep for the pooled data from these three sets, for each speaker.

	initial downstep		downstep after upstep		initial downstep	downstep after upstep	
	H_1	H_2	$\hat{}\underline{\underline{H}}$	H'	$H_2{:}H_1$	$H'{:}\hat{}\underline{\underline{H}}$	t-test
SW	280 (9·3)	248 (10·2)	276 (16·9)	246 (7·9)	0·67 (0·10)	0·69 (0·13)	ns
CB	138 (5·1)	128 (5·5)	133 (4·8)	125 (3·8)	0·71 (0·13)	0·72 (0·14)	ns
TL	167 (10·0)	151 (11·1)	161 (8·5)	137 (6·4)	0·72 (0·13)	0·56 (0·10)	***

Table IV

Comparison of initial downstep with downstep after upstep. The downstep ratios are statistically indistinguishable for speakers SW and CB, while speaker TL shows more dramatic downstep after upstep than initially. Standard deviation is shown between parentheses. All measurements are in Hz.

I will turn to the details of the calculations shortly. For now, notice the main result of the comparison. The statistical comparison confirms the impressions discussed with regard to the plots in Fig. 8 above: speakers SW and CB show downstep relations in medial position that are statistically indistinguishable from the downstep relation in initial position. This provides interesting confirmation for the present analysis, in which upstep is a return to the phrasal reference line, relative to which partial reset is downstepped. Speaker TL shows stronger downstep between phrases than within phrases.[18] This suggests that for some speakers, downstep among i-phrases may be stronger than downstep within i-phrases.

The calculations are based on the theory of Liberman & Pierrehumbert (1984), according to which downstep is treated mathematically as multiplication with a constant fraction, calculated relative to a constant reference line. For a given speaker, then, assuming we know the height of the reference line, we can compare the lowering fraction for downstep in initial position and for downstep after upstep.

The height of the reference line here requires some comment. It is estimated relative to Liberman & Pierrehumbert (1984)'s model, in which the reference line is above the average values of utterance-final L% tones,

[18] A model was fit to the values of TL (taking temporal values of each measurement point into account) to explore whether the low values of H′ may be due to the independently observed declination for speaker TL. However, the angle of declination that can be motivated on the basis of the values of $\hat{}\underline{\underline{H}}$ relative to the initial H_1 is not steep enough to explain the low values in the position of H′.

and below any actual H tones in downstep. The lowering fractions in Table IV are based on a reference line half-way between L % and H_4. The values are intended to give an impression of the downstep ratios and their difference. Importantly, it turns out that the significance results in Table IV for each speaker are independent of choice of reference line, so long as the reference line is chosen between L % and H_4.[19]

The finding that downstep is similar (or stronger) between phrases than within phrases may be compared to the findings of van den Berg *et al.* (1992) for Dutch. The authors found that downstep between phrases in their Dutch data is smaller in extent than downstep within phrases. The present findings suggest that the differences in the Dutch data do not stem from a cross-linguistic strength distinction between downstep within and across phrases.

Summing up the discussion of this section, the consideration of the initial peak in the second clause provides confirmation for the application of the model of van den Berg *et al.* (1992) to the German data at hand in the way suggested – the renewed height initially in the second clause is quite clear, as is the lowering relation between the upstepped (or utterance-initial) level and the beginning of the second clause. Finally, SW and CB show a downstep of the register of the i-phrase that has the same extent as the downstep of the register of the p-phrase. This supports the analysis in which the upstepped tones are scaled to the phrasal register, and in which partial reset is a return to a register that is downstepped relative to the phrasal register of the first i-phrase. TL showed more lowering after the upstepped peak than in initial position, suggesting that downstep on the phrasal level, when occurring on adjacent tones, might sometimes be stronger than downstep within phrases.

3.2.5 *Speaker MG*. Speaker MG differs from the other speakers in a number of ways. In this section, I argue that an independent factor of register undershoot affects the numerical results for this speaker. Once this is taken into account, the combined differences that this speaker shows in fact indirectly support the analysis.

To begin with, the upstepped values in the nuclear pitch accent of MG remain considerably below those of the initial values (initial average 192 Hz, SD 7·8 Hz; upstepped average 181 Hz, SD 6·3 Hz). While the distinction is not large in absolute terms, it is highly significant ($p < 0{\cdot}001$)

[19] The lowering fractions vary with choice of reference line. For SW, for example, the lowering fraction is 0·74 initially and 0·76 medially when calculated against a reference line as low as L %, and 0·52 initially and 0·54 medially when calculated against a reference line as high as H_4. However, the significance results are constant across choice of reference line. For SW, for example, the downstep ratios in initial and medial position are not significantly different from each other when calculated against a reference line of averaged L % values, a reference line of averaged H_4 values or any reference lines between these values, and likewise for CB. For TL, however, the downstep ratios in initial and medial position differ to a highly significant degree ($p < 0{\cdot}001$); that result is again constant for all reference lines between L % and H_4.

and considerable in relative terms – on average, MG executes only 82 % of what would be a full rise to the initial height, and thus considerably less than the speakers discussed above (see the penultimate column of Table II).

Second, there is no complete neutralisation in the position of upstep for speaker MG. This is shown in Table V.[20]

	Means (Hz)				ANOVA		Mean difference (Hz)					
	a	b	c	d	*F*	*df*	a–b	b–c	c–d	a–c	b–d	a–d
MG H	185 (7·0)	177 (4·5)	166 (6·8)	164 (4·8)	45·2 ***	3,68	8·1 ***	9·8 ***	2·4	17·9 ***	12·3 ***	20·3 ***
^H	187 (4·9)	181 (4·4)	179 (5·6)	176 (3·5)	19·5 ***	3,68	6·8 ***	2·1	2·6	8·8 ***	4·7 *	11·4 ***

Table V

Evaluation of the neutralising effect of upstep for speaker MG. Significant distinctions in the immediately prenuclear values (first row) are reflected in significant distinctions in the upstepped nuclear values (second row). Standard deviation (in Hz) is shown between parentheses in the *Means* columns. Significance in the post hoc analysis is given in the *Mean difference* columns.

Table V shows that many of the significant distinctions in the prenuclear H tones lead to significant distinctions in the following ^H values (columns 'a–b', 'a–c', 'b–d' and 'a–d'). Notice, however, that one can nevertheless find a clear neutralising tendency in this data. For one thing, the distinction 'b–c' is completely neutralised. For another, and more importantly, the difference in Hz found in the prenuclear H values is reduced in absolute terms in the nuclear ^H values. Column 'b–d', for example, shows an average difference of 12·3 Hz between prenuclear H after one step of downstep and prenuclear H after three steps of downstep. This leads to an average difference of only 4·7 Hz in the following upstepped values. While the distinction in the upstepped values is still significant, it is smaller in absolute terms than the distinction among the preceding H values. Similarly for the other cases in which the upstepped values retains significant distinctions in the data of MG. MG thus exhibits a clear neutralising tendency in the upstepped values, though no complete neutralisation.

Third, consider the relation between the upstepped values and the following partial reset. As shown in Fig. 9, partial reset is discernible at the beginning of the second clause in the renewed height relative to earlier downstep, and in the relatively lower value of this partial reset than

[20] Tukey's HSD test was applied for the post hoc comparisons of the analysis of variance. The variances are homogeneous by Levene's statistic.

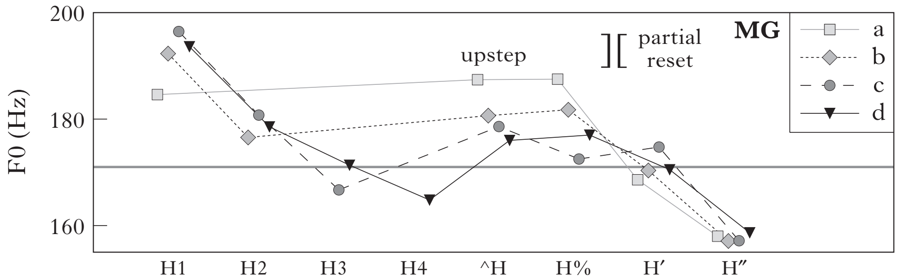


Figure 9

Average values of H tones for speaker MG. The distinctions among sets in the position of upstep (and in H%) do not carry over to the position of partial reset H′.

either the utterance-initial peaks or the upstepped values. However, the relation between upstep and partial reset is unexpected in two ways. For one thing, the relation differs depending on the height of the upstepped values: a relatively larger downstep is found from higher upstepped values, while a very small amount of downstep is found from the lower upstepped values. For another, the downstep relation between the upstepped values and the values in partial reset is on average considerably smaller in extent than the downstep relation in initial position. This can be estimated in Fig. 9, and is confirmed in the statistical evaluation in Table VI. The numbers and significance calculations use a reference line halfway between the lowest downstepped H in the first clause and the final L%.

These differences in the values of MG can be jointly explained if MG shows register undershoot in the H of the upstepped nuclear pitch accent. In other words, we may look upon the phrasal register line as a register target that is not fully reached in the upstepped values of MG. This explains, first, the lower than expected upstepped values, as the target at the height of the utterance-initial peak is not reached, and, second, the partial neutralisation in the upstepped values shown in Table V. In the coarticulation model of Flemming (1997, 2001), in particular, we expect the amount of undershoot of a target to be proportional to the distance of the preceding point from which the target is undershot to the target. In the case at hand, this preceding point is the preceding downstepped register. From there, the amount of undershoot of the higher phrasal register will depend on the distance of the downstepped register to the higher phrasal register: greater distance will lead to larger undershoot. This correctly predicts a distinction in the upstepped values of the kind found here. It also correctly predicts the neutralising tendency in the upstepped values. If the amount of undershoot is proportional to the previous distance from the target, then a difference between two down-

	initial downstep		downstep after upstep		initial downstep	downstep after upstep	
	H_1	H_2	^$\underline{\underline{H}}$	H'	H_2:H_1	H':^$\underline{\underline{H}}$	t-test
MG	194 (6·6)	179 (5·3)	178 (4·9)	172 (5·7)	0·71 (0·09)	0·84 (0·18)	***

Table VI

Comparison of initial downstep with downstep after upstep for speaker MG. There is more dramatic downstep after upstep than in initial position. Standard deviation is shown between parentheses. All measurements are in Hz.

stepped registers is mathematically predicted to lead to a proportionally smaller difference in the two undershot registers. Third and finally, the unexpected downstep relations shown in Fig. 9 and in Table VI can be accounted for. If undershoot is limited to the position of upstep, but the following partial reset is more or less normally executed, then we expect that the downstep relation appears lower in the undershot values than it would be if calculated relative to the target of the phrasal reference line. We also expect the downstep relation to differ from one set to another in the way in which it does, as more undershoot will more dramatically reduce the measured downstep relation.

This interpretation of the data of MG raises a number of interesting questions about the technical implementation of effects of undershoot in a model of phonetic register. However, they are not pursued here, since conclusions we might draw would be based on the data of a single speaker. Implementational questions apart, I believe that looking at MG's data in terms of undershoot is justified by the joint properties of MG's data. I point out, therefore, that MG's data, on this interpretation, fully support the interpretation of upstep defended in this article – interpreting the data as undershoot of a higher target, we have indirect evidence for a higher target, plausibly around the height of the utterance-initial peak, which could serve as a plausible reference point for downstep of the phrasal reference line to the level of partial reset initially in the second clause.

I now sum up the experimental findings. It was seen that the experimental results support the phonetic analysis of upstep in the model of van den Berg *et al.* (1992): as predicted by this analysis, upstep is a return to the utterance-initial height, upstep leads to phonetic values that are independent of preceding downstep, and targets a register relative to which the first peak of the following i-phrase (in partial reset) is downstepped. Where these predictions are not fully borne out, other intervening factors can plausibly be seen to be at play. For speaker TL, this declination was postulated to lead to lower upstepped values, and stronger downstep relation among phrases than within phrases was seen. For speaker MG, a number of observations converge on the conclusion that register undershoot in the position of upstep is at play. Indirectly, however, even the data of this speaker support the phonetic analysis of upstep defended here.

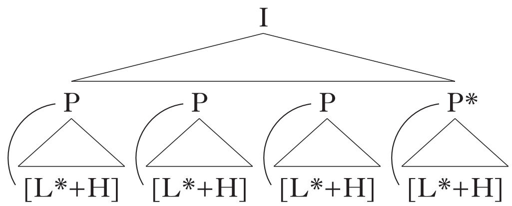
4 Phonological representation

Why, then, may downstep be interrupted on the nuclear pitch accent, and why is the nuclear pitch accent then scaled to the phrasal register? I will offer an answer that has two parts. The first part is a proposal on the phonological representation and on the phonology–phonetics relation that allows for the nuclear accent to be scaled on either the unstepped register of the i-phrase or the downstepped register of the p-phrase. The second part is a suggestion on what may determine the choice between these two options.

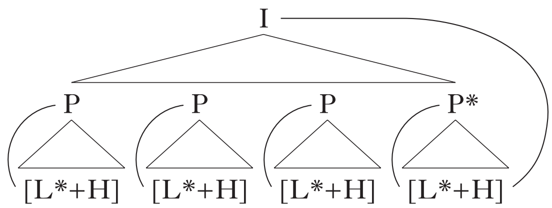
The proposal on the phonology–phonetics relation draws on the suggestions of van den Berg *et al.* (1992) and their implementation of Ladd's idea for a phonetic model in which the upstep can be understood in terms of register levels. It also draws on ideas and suggestions on phonological representations and restrictions in the phonology–phonetics relation from Pierrehumbert & Beckman (1988).

First, I adopt the suggestion of Pierrehumbert & Beckman (1988) that tones in intonation may have direct association to higher prosodic structure. Pierrehumbert & Beckman (1988) formalise and employ this relation for edge tones in Tokyo Japanese. Here I am interested in applying it to pitch accents. Pierrehumbert & Beckman (1988) include a formalisation for associating tones on stressed syllables to higher prosodic constituents as well. They do not apply this in their account of Japanese, but briefly exemplify it with regard to English (1988: 159) – in English, there may be at most one pitch accent on each metrical foot, which is placed on the stressed syllable. One may thus understand accent as a foot-level property. With the pitch accent phonologically associated to the foot, and with this association being interpreted as 'central' (as opposed to 'peripheral', for the edge tones), Pierrehumbert & Beckman's formalisations will entail association to the head of the foot, i.e. the stressed syllable, and thus 'central' placement of the accent within the foot. I suggest that the application of this proposal be expanded. In the German cases discussed above, I suggest that placement of prenuclear pitch accents be formalised as a property of p-phrases (here taken to be the accentual domains). Pierrehumbert & Beckman's conventions will then guarantee inheritance of this association relation from head to head downward to the syllable with strongest stress within the p-phrase, to which the pitch accent is ultimately timed. The nuclear accent now has two options for association. On the one hand, it is the peak of the p-phrase that is final in the i-phrase. Like prenuclear accents, it may thus be associated with the level of the p-phrase, as shown in (9a). In addition, the nuclear accent may be associated at the level of the i-phrase, as in (9b). By the conventions of downward inheritance from head to head, it will then still correctly be placed on the nuclear stress, the strongest stress in the i-phrase. Higher association to the i-phrase, by these conventions, is an option available to the nuclear pitch accent, but not to prenuclear pitch accents.

(9) a.



b.



This difference, I suggest, is the phonological source of the choice of register level. The nuclear pitch accent in (9a) will be scaled in continuation of preceding downstep, while the nuclear pitch accent in (9b) will be scaled to the register of the i-phrase. More generally, association to a particular level is claimed to determine choice of phonetic register. This relation is strengthened in Truckenbrodt (2001), where it is argued to extend to edge tones.

Before formalising this claimed relation between association and scaling, let us once again return to Pierrehumbert & Beckman (1988)'s proposals. Pierrehumbert & Beckman (1988) also suggest multiple simultaneous registers, and correlate them with prosodic levels. They thus distinguish a high-tone line of the intermediate phrase and a high-tone line of the accentual phrase. However, the motivation and nature of this distinction in their account of Tokyo Japanese is quite different from the motivation and nature of the distinction in the account pursued here. The intermediate phrase in Tokyo Japanese is the domain in which reset applies initially, in the terms used here. Pierrehumbert & Beckman conceptualise this by treating the intermediate phrase as the domain that delimits downstep. Consequently, they correlate the register that downsteps within the intermediate phrase as the register pertaining to the prosodic level of the intermediate phrase (in the terms of the present article, this would be the register of the accentual phrase comparable to the p-phrase in German). In Pierrehumbert & Beckman's account, this downstepping register is then distinguished from an additional, lower register (the register of the accentual phrase in their terms), which reflects subordination relations of accentual phrases relative to each other (subordination relations that do not stem from downstep).

The register of the reset domain (intermediate phrase in Pierrehumbert & Beckman's terms), which is imported from the account of van den Berg *et al*. (1992), is an innovation relative to these proposals. Its relation to the reset domain is that it is crucially constant for a given reset domain, thus fulfilling its role in the local computation of non-local downstep in the position of partial reset. With the implementation of this different concept, I propose to change the correlations between prosodic structure and register in a way that follows naturally from van den Berg *et al*.'s model: the register of what I call i-phrase is downstepped once per i-phrase, while the register of the p-phrase is downstepped once per p-phrase (within the same i-phrase). The role of the prosodic categories in defining the course of the register gives an independent basis for the assignment of the phonetic registers to the prosodic levels. This relation is refined in Truckenbrodt (2001).

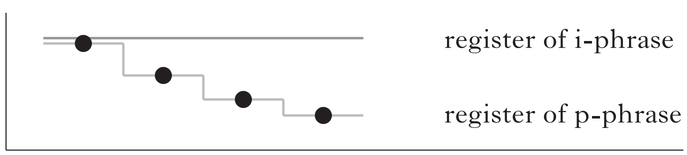
At the same time, I propose to extend the account of Pierrehumbert & Beckman (1988), in which higher prosodic association can condition tonal scaling. The central application of Pierrehumbert & Beckman's idea is that of scaling of a L edge tone of the accentual phrase in Tokyo Japanese. This edge tone will under certain circumstances surface *initially* in an accentual phrase, yet in a crucial case is shown to adhere to the register defined in connection with the *preceding* accentual phrase. Pierrehumbert & Beckman capture this relation of 'backward scaling' by giving L higher association with the preceding accentual phrase. This higher prosodic association is suggested to extend the register defined in the preceding domain to L.

I propose to extend the relation between prosodic association and scaling, as in (10):

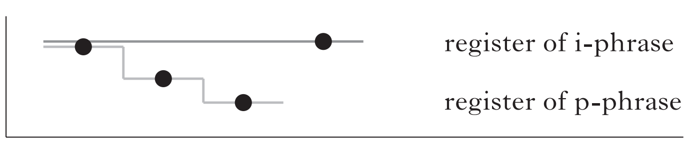
(10) Pitch accents are phonetically scaled to the register that is correlated with the highest prosodic level they are associated with.

By (10), prenuclear pitch accents on p-stress will be scaled to the register of the p-phrase, which is downstepping in the cases considered here. If the nuclear pitch accent is associated no higher than the p-phrase, as in (9a), it will be scaled to the downstepping register of the p-phrase, as in the schematic illustration in (11a). On the other hand, if the nuclear pitch accent is associated with the i-phrase as in (9b), then (10) will lead to the upstepped scaling in (11b).

(11) a.



b.



I turn, then, to the question what may determine low *vs.* high association and scaling of the nuclear pitch accent. I will limit myself to informal speculative remarks, due to limitations of the available data.

An obvious factor that may lead to low association and scaling is UNIFORMITY OF ASSOCIATION: if all pitch accents are associated with the same prosodic level, such as the p-phrase in (11b), low association and scaling in the nucleus results. This may be seen to motivate low association as a general default in English and German. Why, then, do we sometimes find high association and scaling? In the cases of Beckman & Pierrehumbert (1986), Féry (1993) and Fitzpatrick-Cole (1999) reviewed in §2, narrow focus on the element associated with the nucleus was seen to be involved. One may hypothesise that the speaker may choose non-uniform association to set the narrow focus apart from the rest of the utterance. Consider also the case (5b) from Ladd, on the assumption that upstep is here possible without narrow focus, perhaps under emphatic circumstances. In emphasising that a particular point is important, a speaker may generally opt for a phonetically more expanded rendition than the default. While a speaker may do this by generally speaking up, he or she may plausibly be seen to also exploit grammatical means that allow for phonetically higher peaks. In the case at hand, this would be the association of the nucleus to the i-phrase, which allows for a phonetically higher-than-normal nuclear peak.

In my German material, where narrow focus or expressive intonation were deliberately kept out, some speakers, as shown above, chose high association and scaling in utterance-medial position, while all speakers used low association and scaling in utterance-final position. What may determine these choices? In formulating a suggestion, I will call a pitch accent with a H tone, such as L*+H, a 'high pitch accent'. The motivation for associating and scaling high in the nucleus for some speakers may be to have all downstep relations that affect high pitch accents represented in adjacent high pitch accents. As was seen in the Dutch case in (3), with no upstep in the nuclear pitch accent, partial reset is seen on the surface as a downstep at a distance, between the initial high pitch accent and the high pitch accent in partial reset. As was seen in Fig. 5 for the German upstep, the presence of the upstepped nuclear rise makes this abstract downstep relation visible on the surface. If this is the incentive for associating and scaling a pitch accent high, it follows correctly that upstep on the nuclear pitch accent may be found in medial position in my materials, where there is a downstep relation to the following partial reset, but that there is no upstep in utterance-final position, where there is no following partial reset.

5 Summary

I have argued that some speakers of German show an upstep phenomenon on a H tone in the nuclear pitch accent of a non-final intonation phrase, with the following properties: the upstepped height is comparable to the initial height (though declination may affect the comparison), the upstepped peak neutralises distinctions due to preceding downstep, and a partial reset initially in the following clause is phonetically in a downstep relation with the upstepped peak. I have argued that the upstep represents a return to a phrasal register line, as postulated in the model of van den Berg *et al*. (1992). In its relation to the following partial reset, the upstep phenomenon was argued to support the view of the partial reset as downstepped on the phrasal level (Ladd 1988, van den Berg *et al*. 1992). A phonological representation of the upstep is suggested, as is a correlation between prosodic representation and choice of phonetic register. In the phonological representation suggested, the upstepped nuclear pitch accent is associated to the intonation phrase, in an extension of the formalism of Pierrehumbert & Beckman (1988). This is suggested to lead to scaling in the phonetic register of the intonation phrase.

Appendix: the stimuli

a. 1. Die Lola will das Nähen lernen, und die Manu und die Lena sollen ihr eine Nähmaschine kaufen.
   2. Die Hanne soll Bananen essen, und die Lena und der Werner sollen ihr welche mitbringen.
   3. Der Maler will Nonnen malen, und die Lena und die Hanne sollen ihm Modell stehen.
   4. Der Maurer will das Weben lernen, und die Hanne soll ihm Leinen und Wolle besorgen.
   5. Die Lena will eine Malve malen, und im Januar will der Werner nach Hamburg gehen.
   6. Der Werner soll Romme lernen, und der Rommelehrer soll aus Lüneburg oder aus Hamburg kommen.

b. 1. Der Werner und die Lola wollen malen lernen, und die Manu will dem Lehrer Romme zeigen.
   2. Der Lehrer und der Maurer wollen nähen lernen, und die Nonne will dem Heiner Wolle leihen.
   3. Die Nonne und der Lehrer wollen malen lernen, und die Lena will dem Werner einen Roman geben.
   4. Die Lena und die Manu wollen die Nonne malen, und der Werner soll in Murnau einen Roman schreiben.
   5. Der Lehrer und die Hanne wollen nähen lernen, und der Maler und die Lena wollen nach Hamburg ziehen.
   6. Der Wladimir und die Hanne wollen ein Lama malen, und die Lola soll der Lena Wolle leihen.

c. 1. Ein Tiger und ein Löwe wollen einem Lama Maronen wegnehmen, und ein Marabu und eine Möwe wollen einen Wurm fangen.
   2. Die Lena und die Hanne wollen einem Maurer das Weben beibringen, und der Manuel will die Lola in Murnau besuchen.
   3. Der Maler und der Lehrer wollen der Hanne Maronen geben, und der Maurer will der Lena Murnau zeigen.
   4. Der Werner und die Lena wollen dem Lehrling Manieren beibringen, und die Lola will dem Manuel eine Warnung geben.

5. Der Werner und die Lena wollen der Nonne ein Lama malen, und der Heiner will in Hamburg eine Mole mauern.
6. Der Maurer und die Nonne wollen dem Lehrer Malvenrum bringen, und der Werner will der Manu einen Roman vorlesen.

d. 1. Die Nonne und der Lehrer wollen der Lola in Murnau eine Warnung geben, und die Hanne will im November ein Lama malen.
2. Die Lola und die Manu wollen der Lena im November Maronen geben, und der Werner will der Hanne ihr Leinen weben.
3. Die Lena und die Lola wollen dem Werner im Januar ein Lama malen, und die Manu soll im November Bananen essen.
4. Die Manu und die Hanne sollen der Lena im Januar das Leinen weben, und der Werner soll in Murnau Maronen holen.
5. Der Maler und der Lehrer wollen der Lena in Malmö einen Marabu geben, und die Lola will sich im Januar eine Wohnung nehmen.
6. Der Maurer und sein Lehrling wollen dem Werner in Kamerun ein Lama malen, und der Maler will im Januar in Murnau wohnen.

The preceding stimuli were used for all speakers except SW, with whom minimally different stimuli were used in earlier recordings. These were later made more uniform for the other speakers. The differences in the stimuli for SW are as follows.

The (a) series had shorter second clauses (see note 17), but identical first clauses; this was later changed in the stimuli for the other speakers, for reasons of uniformity. The shorter second clauses used with SW are:

a. 1. … , und die Manu soll ihr dabei helfen.
2. … , und die Lena will ihr welche kaufen.
3. … , und die Lena soll ihm Modell stehen.
4. … , und die Lehrerin soll ihm dabei helfen.
5. … , und der Werner will lieber rumgammeln.
6. … , und der Rommelehrer soll aus Hamburg kommen.

In the other series, longer words of the stimuli with SW were exchanged for shorter words in the stimuli used for all others. The sentences in which these were substituted otherwise remained the same. These were: (c2), (c3), (d1), (d2) (*Hannelore/Hanne*); (c6) (*Rommelehrer/Lehrer*); (c4) (*Maurerlehrling/Lehrling*). In the (b) series, three sentences had longer words more systematically, and were replaced with different sentences altogether. The sentences with longer words (with an analogous prosodic pattern in the relevant respects) used for SW are:

b. 4. Kamerun und die Niederlande sollen Lehrermangel haben, und Panama will Norwegen ein Walfang-Abkommen anbieten.
5. Der Rommelehrer und die Hannelore wollen in einer Lagerhalle wohnen, und der Maler und die Lena wollen nach Hamburg ziehen.
6. Der Wladimir und die Hannelore wollen Mengenlehre lernen, und die Lola soll der Lena ein Lama malen.

The substitution of shorter words was made for reasons of uniformity. The longer words were otherwise both expected and found to participate in a regular fashion in the tonal patterns also elicited with the shorter words.

Finally, SW used the dialectal *Jänner* instead of *Januar*.

REFERENCES

Altmann, Hans (ed.) (1988). *Intonationsforschungen*. Tübingen: Niemeyer.

Altmann, Hans, Anton Batliner & Wilhelm Oppenrieder (eds.) (1989). *Zur Intonation von Modus und Fokus im Deutschen*. Tübingen: Niemeyer.

Arvaniti, Amalia, D. Robert Ladd & Ineke Mennen (2000). What is a starred tone? Evidence from Greek. In Broe & Pierrehumbert (2000). 119–131.

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. 335–367.

Bierwisch, Manfred (1966). Regeln für die Intonation deutscher Sätze. *Studia Grammatica* **7**. 99–201.

Broe, Michael B. & Janet B. Pierrehumbert (eds.) (2000). *Papers in laboratory phonology V: acquisition and the lexicon*. Cambridge: Cambridge University Press.

Bruce, Gösta (1977). *Swedish word accents in sentences perspective*. Lund: Gleerup.

Bruce, Gösta & Eva Gårding (1978). A prosodic typology for Swedish dialects. In Eva Gårding, Gösta Bruce & Robert Bannert (eds.) *Nordic prosody*. Lund: Department of Linguistics, Lund University. 219–228

Büring, Daniel (1995). *The 59th Street Bridge accent: on the meaning of topic and focus*. PhD dissertation, University of Tübingen.

Büring, Daniel (2001). What do definites do that indefinites definitively don't? In Caroline Féry & Wolfgang Sternefeld (eds.) *Audiatur vox sapientiae: a Festschrift for Arnim von Stechow*. Berlin: Akademie Verlag. 70–100.

Cinque, Guglielmo (1993). A null theory of phrase and compound stress. *LI* **24**. 239–297.

Clements, George N. (1979). The description of terraced-level tone languages. *Lg* **55**. 536–558.

de Lacy, Paul (1998). Higher prosody in Maori. Ms, University of Massachusetts, Amherst.

Diesing, Molly (1992). *Indefinites*. Cambridge, Mass: MIT Press.

Essen, Otto von (1964). *Grundzüge der hochdeutschen Satzintonation*. Ratingen: Henn.

Féry, Caroline (1992). Focus, topic and intonation in German. *Arbeitspapiere des Sonderforschungsbereichs 340*, '*Sprachtheoretische Grundlagen für die Computerlinguistik*' **20**, University of Tübingen.

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

Féry, Caroline (1995). *Alignment, syllable and metrical structure in German*. Habilitationsschrift, University of Tübingen.

Fitzpatrick-Cole, Jennifer (1999). The Alpine intonation of Bern Swiss German. Paper presented at 14th International Congress of Phonetic Sciences, San Francisco.

Flemming, Edward (1997). Phonetic optimization: compromise in speech production. In Viola Miglio & Bruce Morén (eds.) *University of Maryland Working Papers in Linguistics* **5**. *Selected Phonology Papers from the Hopkins Optimality Theory Workshop 1997/University of Maryland Mayfest 1997*. 72–91.

Flemming, Edward (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* **18**. 7–44.

Goldsmith, John (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, D. Robert Ladd & Amalia Arvaniti (2000). On the place of phrase accents in intonational phonology. *Phonology* **17**. 143–185.

Grice, Martine, Stefan Baumann & Ralf Benzmuller (to appear). German intonation within the framework of autosegmental-metrical phonology. In Sun-Ah Jun (ed.) *Prosodic typology and transcription : a unified approach*. Oxford: Oxford University Press.
Gussenhoven, Carlos (1983). Focus, mode and the nucleus. *JL* **19**. 377–417.
Gussenhoven, Carlos (1984). *On the grammar and semantics of sentence accents*. Dordrecht: Foris.
Gussenhoven, Carlos (1992). Sentence accents and argument structure. In Iggy M. Roca (ed.) *Thematic structure : its role in grammar*. Berlin & New York: Foris. 79–106.
Gussenhoven, Carlos (2000a). The boundary tones are coming: on the nonperipheral realization of boundary tones. In Broe & Pierrehumbert (2000). 132–151.
Gussenhoven, Carlos (2000b). The lexical tone contrast of Roermond Dutch in Optimality Theory. In Horne (2000). 129–167.
Gussenhoven, Carlos & Toni Rietveld (1988). Fundamental frequency declination in Dutch: testing three hypotheses. *JPh* **16**. 355–369.
Gussenhoven, Carlos & Toni Rietveld (1998). On the speaker-dependence of the perceived prominence of F_0 peaks. *JPh* **26**. 371–380.
Gussenhoven, Carlos, Bruno H. Repp, Toni Rietveld, H. H. Rump & Jacques Terken (1997). The perceptual prominence of fundamental frequency peaks. *JASA* **102**. 3009–3022.
Hayes, Bruce & Aditi Lahiri (1991). Bengali intonational phonology. *NLLT* **9**. 47–96.
Höhle, Tilman N. (1982). Explikationen für 'normale Betonung' und 'normale Wortstellung'. In Werner Abraham (ed.) *Satzglieder im Deutschen : Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung*. Tübingen: Narr. 75–153.
Horne, Merle (ed.) (2000). *Prosody : theory and experiment*. Dordrecht: Kluwer.
Isačenko, Alexander V. & Hans-Joachim Schädlich (1966). Untersuchungen über die deutsche Satzintonation. *Studia Grammatica* **7**. 7–67.
Jacobs, Joachim (1993). Integration. In Marga Reis (ed.) *Wortstellung und Informationsstruktur*. Tübingen: Niemeyer. 63–116.
Jäger, Gerhard (1995). Topic, scrambling und aktionsart. *Arbeitspapiere des Sonderforschungsbereichs 340*, '*Sprachtheoretische Grundlagen für die Computerlinguistik*' **69**, University of Tübingen. 19–34.
Kingston, John & Mary E. Beckman (eds.) (1990). *Papers in laboratory phonology I : between the grammar and physics of speech*. Cambridge: Cambridge University Press.
Kohler, Klaus J. (ed.) (1991). *Studies in German intonation*. Kiel: Institut für Phonetik und digitale Sprachverarbeitung, University of Kiel.
Krifka, Manfred (1984). *Focus, Topic, syntaktische Struktur und semantische Interpretation*. Ms, University of Munich.
Krifka, Manfred (1998). Scope inversion under the rise-fall contour in German. *LI* **29**. 75–112.
Kubozono, Haruo (1989). Syntactic and rhythmic effects on downstep in Japanese. *Phonology* **6**. 39–67.
Kubozono, Haruo (1993). *The organization of Japanese prosody*. Tokyo: Kurosio.
Ladd, D. Robert (1983a). Phonological features of intonational peaks. *Lg* **59**. 721–759.
Ladd, D. Robert (1983b). Even, focus, and normal stress. *Journal of Semantics* **2**. 157–170.
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 Kingston & Beckman (1990). 35–57.
Ladd, D. Robert (1993a). On the theoretical status of 'the baseline' in modeling intonation. *Language and Speech* **36**. 435–451.
Ladd, D. Robert (1994). Constraints on the gradient variability of pitch range, or,

Pitch level 4 lives! In Patricia Keating (ed.) *Phonological structure and phonetic form: papers in laboratory phonology III*. Cambridge: Cambridge University Press. 43–63.
Ladd, D. Robert (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
Ladd, D. Robert, Ineke Mennen & Astrid Schepman (2000). Phonological conditioning of peak alignment in rising pitch accents in Dutch. *JASA* **107**. 2685–2696.
Ladd, D. Robert & Rachel Morton (1997). The perception of intonational emphasis: continuous or categorical? *JPh* **25**. 313–342.
Ladd, D. Robert, Jo Verhoeven & Karen Jacobs (1994). Influence of adjacent pitch accents on each other's perceived prominence: two contradictory effects. *JPh* **22**. 87–99.
Laniran, Yetunde (1992). *Intonation in tone languages: the phonetic implementation of tones in Yorùbá*. PhD dissertation, Cornell University.
Liberman, Mark & Janet 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.
Myers, Scott (1999). Tone association and f_0 timing in Chichewa. *Studies in African Linguistics* **28**. 215–239.
Nespor, Marina & Irene Vogel (1986). *Prosodic phonology*. Dordrecht: Foris.
Odden, David (1995). Tone: African languages. In Goldsmith (1995). 444–475.
Pheby, John (1984). Phonologie: Intonation. In Karl Erich Heidolph, Walter Flämig & Wolfgang Motsch (eds.) *Grundzüge einer deutschen Grammatik*. Berlin: Akademie Verlag. 839–897.
Pierrehumbert, Janet B. (1980). *The phonology and phonetics of English intonation*. PhD dissertation, MIT.
Pierrehumbert, Janet B. (1994). Alignment and prosodic heads. *Proceedings of the Eastern States Conference on Linguistics* **10**. 268–286.
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.
Prieto, Pilar, Jan van Santen & Julia Hirschberg (1995). Tonal alignment patterns in Spanish. *JPh* **23**. 429–451.
Prieto, Pilar, Chilin Shih & Holly Nibert (1996). Pitch downtrends in Spanish. *JPh* **24**. 445–473.
Prince, Alan & Paul Smolensky (1993). *Optimality Theory: constraint interaction in generative grammar*. Ms, Rutgers University & University of Colorado, Boulder.
Rooth, Mats (1992). A theory of focus interpretation. *Natural Language Semantics* **1**. 75–116.
Schmerling, Susan (1976). *Aspects of English sentence stress*. Austin: University of Texas Press.
Schwarzschild, Roger (1999). Givenness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics* **7**. 141–177.
Selkirk, Elisabeth (1984). *Phonology and syntax: the relation between sound and structure*. Cambridge, Mass.: MIT Press.
Selkirk, Elisabeth (1995). Sentence prosody: intonation, stress, and phrasing. In Goldsmith (1995). 550–569.
Selkirk, Elisabeth & Koichi Tateishi (1991). Syntax and downstep in Japanese. In Carol Georgopolous & Roberta Ishihara (eds.) *Interdisciplinary approaches to language: essays in honor of S.-Y. Kuroda*. Dordrecht: Kluwer. 519–543.
Silverman, Kim E. A. & Janet B. Pierrehumbert (1990). The timing of prenuclear high accents in English. In Kingston & Beckman (1990). 72–106.

Stechow, Arnim von & Susanne Uhmann (1986). Some remarks on focus projection. In Werner Abraham & Sjaak de Meij (eds.) *Topic, focus, and configurationality*. Amsterdam: Benjamins. 295–320.
Terken, Jacques (1991). Fundamental frequency and perceived prominence of accented syllables. *JASA* **89**. 1768–1776.
Terken, Jacques (1993). Baseline revisited: reply to Ladd. *Language and Speech* **36**. 453–459.
Terken, Jacques & Dik Hermes (2000). The perception of prosodic prominence. In Horne (2000). 89–127.
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 (2001). On the scaling of high edge tones in German. Ms, Rutgers University.
Uhmann, Susanne (1991). *Fokusphonologie*. Tübingen: Niemeyer.
Vennemann, Theo (1982). Zur Silbenstruktur der deutschen Standardsprache. In Theo Vennemann (ed.) *Silben, Segmente, Akzente*. Tübingen: Niemeyer. 261–305.
Wiese, Richard (1996). *The phonology of German*. Oxford: Clarendon.
Wunderlich, Dieter (1988). Der Ton macht die Melodie: zur Phonologie der Intonation des Deutschen. In Altmann (1988). 1–40.
Xu, Yi (1998). Consistency of tone–syllable alignment across different syllable structures and speaking rates. *Phonetica* **55**. 179–203.