The variability of early accent peaks in Standard German

1. Introduction

The autosegmental-metrical (A-M) theory of intonation rests on the premise that an f0-contour is derived from a sequence of abstract phonological tone-targets that are systematically aligned with segmental units (Beckman & Pierrehumbert, 1986; Ladd, 1996; Pierrehumbert, 1980). The inventory is reduced to a basic distinction between H(igh) and L(ow) tones and a starred notation is used in pitch-accents to indicate the phonological association between a tone and a rhythmically strong syllable which in turns has a major influence on their temporal alignment. In one sense, temporal alignment is clearly phonological because it is the result of the way that the tune is associated with the text, but in another alignment is also context-dependent and phonetic as the various experiments by Pierrehumbert and Silverman (1990) have shown. The frequency scaling of the High and Low tones is assumed to be predictable from phonetic factors and controlled by a system of rules. As Liberman and Pierrehumbert (1984) have argued, downstep is one of the mechanisms used to justify the adherence to a two-tone model in the face of what are evidently many different f0-levels on the surface. Thus, in the famous stepping contour, the progressive lowering of f0-levels on the strong syllables of successive accented words can be modelled by downstep that is iteratively (and in theory infinitely) applied to a H* tone within a single phrase. This specification of downstep is clearly phonetic.

However, it has often been difficult to place 'downstep' along the divide between phonological categories and phonetic implementation. Downstep was originally invoked in tone languages to express the idea that there is a High tone that is automatically lowered after a low tone in an H-L-H sequence and in the A-M model, it refers to the lowering of successive High tone targets within a prosodic phrase. Unlike Liberman & Pierrehumbert (1984), Ladd (1996: 97) argues that there is also evidence for downstep to be largely an 'orthogonal phonological variable' with the meaning of 'finality' or 'completeness' (cf. Ladd, 1996: 90-92; this was also argued in Ladd, 1983).

The downstepped pitch-accents in American English include monotonal !H* and bitonal L+!H*. Since they necessarily follow a non-downstepped and often bitonal accent, these downstepped accents cannot occur in the phrase-initial position. Our interest lies in a third type of downstepped tone H+!H* which arose in the evolution of the ToBI system, based on the A-M model, for transcribing standard American English intonation. In the ToBI system, H+L* came to be replaced by H+!H* (Ladd 1996: 96-97; Gussenhoven 2004: 132) both for phonetic reasons and because the level of the starred tone was not usually at the bottom of the speaker's range, as would be expected if the second component of this bitonal accent really were L*. Given that one tone essentially replaces another in the ToBI system, there is obviously no sense in which they can be phonologically contrastive in a ToBI transcription of American English. However, the research on intonational phonology by Grice (1995) in Southern British English and by Grice & Baumann (2002) in standard German does provide evidence that there may indeed be a contrast between H+L* and H+!H* and it is this aspect that we wish to explore for German using acoustic and perceptual techniques.

Some examples of this contrast are presented in the German GToBI training materials (Benzmüller, Grice, and Baumann, URL1). In both H+L* and H+!H*, there is a pitch peak due to the H+ that precedes the accented vowel: the difference between these pitch accents is that in H+L* a *low pitch target* is reached in the accented vowel, whereas in H+!H* there is a *downstepped peak* that occurs during the accented vowel, so that the target step is from high to mid in H+!H* but not in H+L*. This is further illustrated in Fig. 1 which shows a H+L* pitch-accent on 'schräg' ('diagonally') in the left panel and a H+!H* pitch-accent on 'schön' in the right panel. There is a marked step down in pitch from the preceding syllable to a pitch-trough that occurs during the [e] of 'schräg'; on the other hand, there is no evidence of a trough in the [ø] vowel of 'schön' and the pitch does not reach the bottom of the speaker's range as it does for H+L* in 'schräg'.

If we allow the system to contrast H+L* and H+!H* pitch-accents, then this means that the frequency scaling of the (starred) tones that are associated with the accented syllable is no longer phonetic as argued in e.g., Pierrehumbert (1980), but is implicitly phonological because there is now a paradigmatic contrast between *three* tonal levels: high-star in (L+)H*, low-star in H+L* and effectively a mid-star in H+!H*. Furthermore, H+!H* represents a departure from the phonotactics of downstep discussed earlier, since there is no obligation for the downstepped H+!H* to follow a non-downstepped pitch-accent (as a result of which, H+!H*, unlike !H* or L+!H*, could be the first and indeed only pitch-accent in an intermediate phrase).

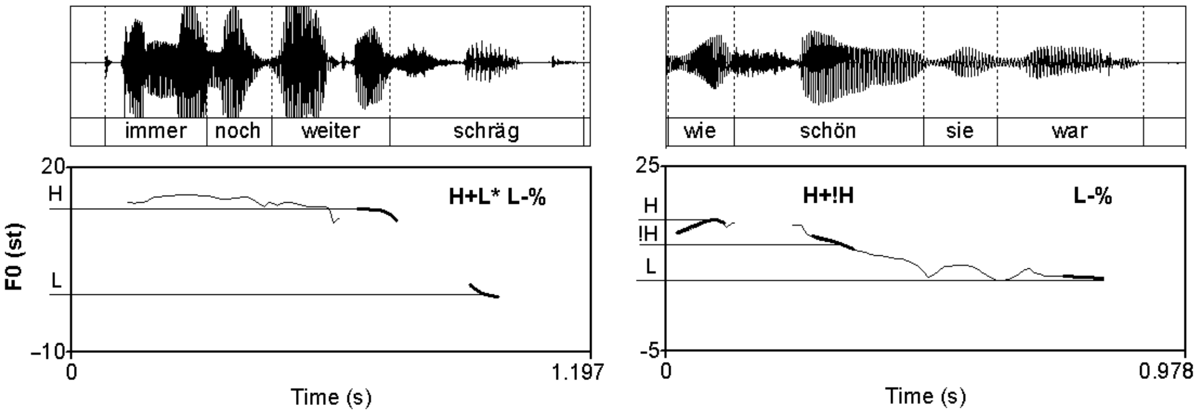

Figure 1: Synchronized time-waveform with word labels (top) and F0 trajectories with GToBI labels of nuclear accents (bottom) taken from Benzmüller, Grice, and Baumann (URL 1) of two German utterances: *"…immer noch weiter <u>schräg</u>."* ('still further along diagonally ') and *"…wie <u>schön</u> sie war."* ('how pretty it was').

There is some difficulty in establishing the meaning differences associated with these pitch accents. Baumann (pers. comm.) suggests that H+!H* might convey a reassuring, soothing tone of voice in contrast to the greater matter-of-factness and finality of a H+L*. In a previous study (Rathcke and Harrington, in press), we assumed the general semantic difference to be that between general/polite (H+!H*) and resolute/self-evident (H+L*) statements, but we were not able to verify any of these hypothesised meaning-tune associations has been verified empirically.

In other models of the intonation of standard German (e.g., Grabe, 1998; Féry 1993; Kohler 1997), there is clear evidence for a contrast between H* and H+L*, i.e. for a contrast of timing in which the pitch-peak is synchronised near the temporal midpoint of the vowel in H* as opposed to its much earlier synchronisation in H+L*, but there have been no suggestions to our knowledge that there is a three-way phonological contrast between H* and what would be two types of early peaks, H+L* and H+!H*. We therefore have to consider the possibility that the distinction between H+L* and H+!H* may not be phonological but phonetic, i.e. predictably related to factors such as phrasal position, number of syllables, of speech rate.

Our basis for assessing whether the distinction between two pitch-accents is categorical is firstly that native listeners can hear the distinction and secondly that it is associated with semantic differences. Since, for the reasons outlined above, it is difficult to be certain about the type of meaning difference that is likely to be associated with the pitch-accent distinctions being investigated, the present paper draws upon a semantic differential test for this purpose (Ambrazaitis 2005; Dombrowski 2003; Osgood, Suci, and Tannenbaum 1957; Uldall 1964). In this test, listeners are asked to make judgements on various semantic scales. We also carried out an imitation test of the kind used in Pierrehumbert and Steele (1989) in which subjects imitate tokens from a continuum spanning the tonal categories under investigation.

We created a synthetic continuum in the f0-scaling domain spanning three pitch-accent categories, H*, H+!H* and H+L* with the H* end of the continuum as control stimuli. We chose to make our continuum span three rather than just two (H+!H*, H+L*) categories, firstly because we do not know at which f0-level over the pitch-accented syllable the continuum should begin for a H+!H* accent (Rathcke and Harrington, in press); and secondly because synthesising downstep would need a speaker-independent downstep-coefficient (van den Berg, Gussenhoven, und Rietveld, 1992) and it is unclear from the literature what value this coefficient should have.

If the distinction between H+!H* and H+L* is categorical in the same way as has been previously shown for the distinction between H* and H+L*, then listeners should be more likely to perceive, and to produce, between- rather than within-category differences: that is, subjects' responses to the synthetic continuum should be trimodal. On the other hand, subjects' responses will be bimodal due to a contrast between H* vs. a collapsed H+!H*/H+L* category, if the latter are just phonetic variants of a single 'early peak' category. Similarly, we would expect three different types of responses to the continuum in the semantic differential test if all three pitch-accents are phonologically distinct.

2. Experiment

2.1. Stimuli

The test sentence *'Sie mag Bananen'* ('she likes bananas') was read several times by a female speaker as a simple statement with a nuclear accent on the medial rhythmically strong syllable *'-na-'* in *'Bananen'*. The production in which she produced the final weak syllable *'-nen'* with a schwa and in which the pitch-accent was clearly H+L* was chosen as the basic item for resynthesis purposes. This was done for primarily because our pilot study (Rathcke and Harrington, in press) had shown that tokens of 'Bananen' with no final schwa vowel were often imitated with a creaky voice extending from some part of the accented syllable of 'Bananen to its offset, thereby making f0 difficult to measure.

The construction of the continuum was based on the idea, proposed by Grice and Baumann (2002: 278), that the main differences between H*, H+!H*, and H+L* are in the *f0-height of accented vowel*. The continuum was therefore designed to extend over three categories from H* through H+!H* to H+L* by incrementing f0-height variations in the accented vowel. More specifically, we created a stylised H+L* contour from the subject's actual production of H+L* by making the f0-contour fall over the accented syllable /'na/ in a straight line from the onset of the initial [n] at 210 Hz to 140 Hz at a point 1/3 of the way into the vowel. For this stylised H+L*, the f0 was kept level at 140 Hz to the end of [a] as shown in Fig. 2 (st_{16}). Two points (at the onset and offset of the accented syllable) remained fixed in time and frequency for all synthesised tokens. The continuum designed to span the differences between the three pitch-accents was synthesised by raising the f0-level in fifteen 0.5 semitone steps over the middle 1/3 section of the vowel and then by connecting the beginning and end of this raised section with straight lines as shown in Fig. 2.

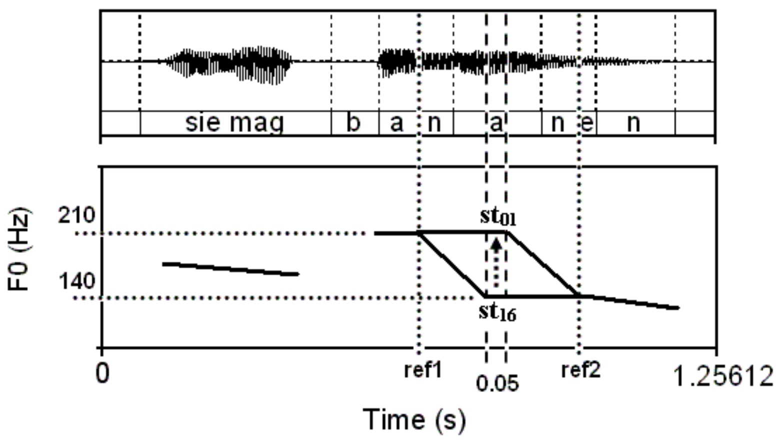

Figure 3: Time-waveform with word and segment labels of the test sentence (top) synchronized with F0 trajectories of the first (st_{01}) and the last (st_{16}) stimulus of the continuum. For the all tokens from the continuum, the time and frequency points marked by *ref1* and *ref2* on the time axis were fixed. The vertical dashed lines mark the interval over the middle 1/3 section of the vowel that was raised in sixteen 0.5 semitone steps.

2.2. Semantic scales

The semantic differential measures subjects' reactions to stimuli in terms of ratings on bipolar scales typically defined with contrasting adjectives at each end. The list of semantic scales resulting in a semantic differential is primarily chosen based on hypotheses about the meaning of the categories under investigation. As has been shown in a various studies (see Heise, 1970, for an overview), ratings on bipolar scales tend to be correlated, and the three basic dimensions of response (labelled 'evaluation', 'potency' and 'activity') account for most of the co-variation in ratings. Using a few pure scales like 'good-bad' for evaluation, 'powerful-powerless' for potency, and 'fast-slow' for activity, enable reliable measures to be obtained of subjects' overall attitudinal responses to different stimuli.

In contrast to the earlier semantic differential studies (e.g. Uldall 1964) which were characterised by an *a posteriori* factor analysis for the purposes of extracting the underlying semantic dimensions, a complete semantic differential including the basic semantic dimensions was constructed *a priori* (see e.g., Ambrazaitis 2005 and Dombrowski 2003 for more details). For the present investigation, eight scales were created. The choice of scales was based on the following considerations:

1. Scales 1-3: These correspond to meaning contrasts that have already been demonstrated for distinctions analgous to H* vs. H+L* (e.g., Kohler, 1987; 2005)
2. Scale 4: This was chosen from considerations that are not unrelated to the frequency code (Ohala 1984): perhaps H+!H* is used paralinguistically to avoid very low f0-values (which are assumed to be typically for H+L* pitch-accents). As is well known (e.g. Gussenhoven, 2002), low tones tend to convey speakers' dominance or to signal unalterable facts in a great number of languages.
3. Scale 5: As outlined in the introduction, this scale's semantic distinction has already been hypothesised as one of the ways that meaning differences between H+!H* and H+L* may be conveyed.
4. Scales 6-8: These were added to complete the semantic differential with the three basis dimensions (evaluation, potency, activity, respectively). A relevant factor for the choice of contrasting adjectives for the scale 6 is that Grice & Baumann (2002) consider 'politeness' to be one of the meanings associated with the H+!H* vs. H+L* distinction. The semantic contrasts chosen for the scales 7 and 8 were thought to be good exponents for 'potency' and 'activity' in dialogues.

The scales are shown in Table 1. Taking into account the meaning distinctions associated with the frequency code (Ohala 1984) and what is known about how meaning differences are conveyed by the distinctions between H*, H+!H* and H+L* (Grice and Baumann, 2002; Kohler, 1987 and 2005), the negative pole of each scale was assumed to be more appropriate for H+L* pitch-accents, the positive pole in scales 1-3 for H* and the positive pole of scales 4-8 for H+!H* pitch-accents.

The scales were embedded in a carrier sentence *"The speaker sounds like…"* and the instructions to the subjects for judging the scalar values made it clear that we were interested in potential meaning differences evoked by the speaker's melody.

Table 1: Eight pragmatic scales chosen for the perception test with semantic differential

Scale no.	Negative pole (-)	Positive pole (+)
1	knowing	realising
2	concluding	opening a discussion
3	accepting	refusing
4	out of the question	ready to compromise
5	calming	encouraging
6	impolite	polite
7	certain	uncertain
8	bored	interested

2.3. Procedures

A tape containing 128 stimuli was created from eight repetitions of each of the 16 stimuli. Thus each stimulus was paired once with each scale and presented in randomised order. The tape was played to the subjects who had to mark their answers on a sheet of paper containing the scales, whether the stimulus was most appropriate for the one or the other end of the given scale and to which extent. Seven point scales were given for decisions with the position 0 labeled "neutral", the 1 positions labeled "slightly", the 2 positions "quite", and the 3 positions "extremely".

The subjects were instructed to wait for the auditory prompt preceded by a beep and to judge spontaneously on the given scale how they thought the speaker sounded. They did this for each stimulus separated by a 4 s pause. The perception task was preceded by a trial session including eight stimuli presentations (one for each scale) to familiarise the subjects with scales and procedure. After the trial session the subjects had the opportunity to ask questions. The stimuli were presented from a CD-player in a sound treated room at the IPDS Kiel.

For the imitation test, the sixteen stimuli detailed above were each copied ten times resulting in 160 items. These 160 items were randomised and then each presented twice with a preceding beep. After the second presentation of each stimulus, there was a pause during which the subject was instructed to imitate it paying particular attention to copying the melody as closely as possible. In the event of a hesitation or speech error, the item was repeated. No time limit was imposed for responses. The stimuli were presented to subjects via headphones. The imitation test was carried out in a sound treated recording studio at the IPDS Kiel.

Twelve speakers of standard North German, five M and seven F of between 20 and 22 years of age with no known speech or hearing disorders participated in the experiment. All subjects were beginner students of phonetics at the IPDS Kiel. None of the subjects had any experience in prosody nor was told of the purpose of the experiment. All subjects were paid for participation.

2.4. Results

2.4.1. Semantic Differential

The evaluation of the 16-point continuum is shown in Fig. 3. The data for each scale were analysed separately by a repeated measures analysis of variance using SPSS software (eight ANOVAs, s. Brosius, 2002). The f0-level in the accented vowel (i.e., stimulus number) was the independent variable and the judgement score the dependent variable; we used a repeated measures ANOVA because the subjects judged more than one stimulus within one analysis. The Greenhouse-Geisser correction for non-homogenity of variances was applied (Leonhart, 2004). These statistical results are summarised in Table 2.

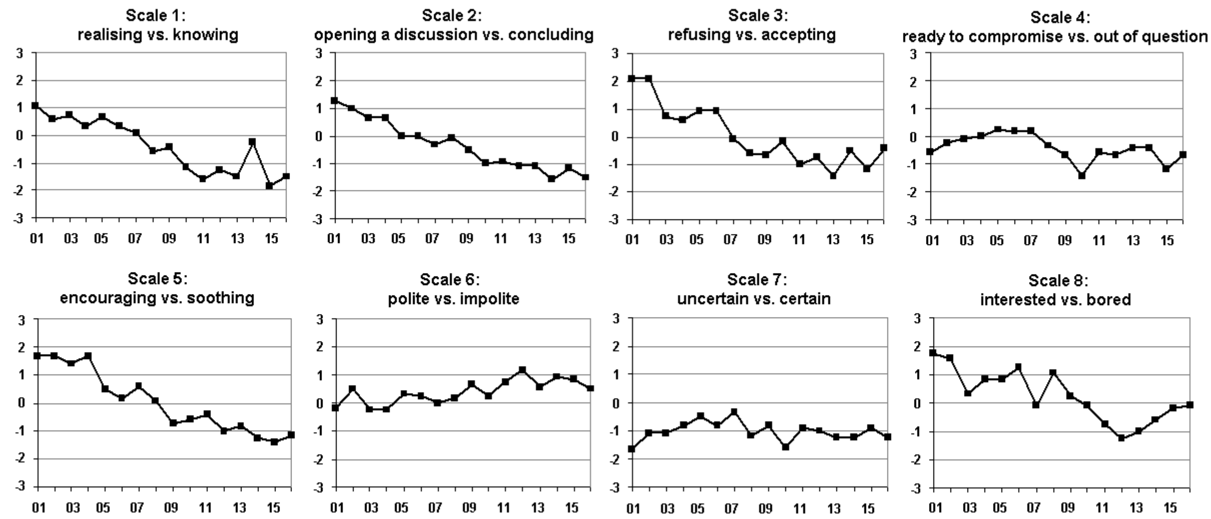

Figure 3: Mean judgements (n=12) for the stimuli st_{01}-st_{16} on the eight scales: x-axis is stimulus number, y-axis is the scale value. The left term above each scale is associated with positive scale values.

Table 2: Results of repeated measures ANOVAs examining the effect of eight scales on the perception of a meaning change during the tonal continuum (Alpha = 1%): F values, adjusted degrees of freedom (d.f.) and significance levels.

	F	**df**	**p**
Scale 1	4.855	5.434	< 0.001
Scale 2	4.319	5.423	<0.01
Scale 3	7.414	5.078	< 0.001
Scale 4	1.082	3.664	n.s.
Scale 5	7.389	5.706	< 0.001
Scale 6	1.179	6.219	n.s.
Scale 7	0.757	4.934	n.s.
Scale 8	6.544	5.346	< 0.001

There were significant differences in subject responses on 5/8 of the scales. Scales 1-3 showed changes of judgement that are compatible with the semantic differences for the distinction between 'medial' and 'early' peaks (Kohler, 1987 and 2005), corresponding roughly to the H* vs. H+L* difference. Scale 5 shows similar results to scales 1-3, i.e. the meaning ‘soothing’ show differences primarily between the beginning and end of the scales, and hence between H* and H+L* and not – as hypothesised – between H+!H* and H+L* (in which case, we would expect a major difference between the middle and the end of the scale). The same is true for the scale 8 although the profile of judgments here is not as marked as for scales 1-3 and 5.

All stimuli were interpreted as neutral on the scales 4 and 6. There were also no perceived differences on the scale 7, i.e. all stimuli of the continuum were judged to be 'certain statements'.

In summary, the first 4-5 points of the stimuli seem to be compatible with the semantic interpretation of a H* pitch-accent, the last 7 stimuli are compatible with the semantics of an early peak, while three stimuli from the middle of the continuum (st_{06}-st_{08}) mark a transition in which there is a strong variation in semantic judgements. There is no evidence for a category formation in the middle of the synthetic continuum which could be labelled H+!H*.

2.4.2. Imitation

The data from 2 subjects were removed from the evaluation: one subject had creaky voice and the other produced all the stimuli on a monotone. We measured the difference (in semitones) between the first two /a/ vowels of /ba'na/ in the subjects' imitations of the synthetic stimuli. Recall that the relationship between the synthetic stimuli and the tonal categories is as follows. In H+L* and H+!H*, f0 is expected to fall across the sequence of weak-strong syllables in /ba'na/ but that the fall is expected to be greater in H+L*. In H* by contrast, the pitch is either likely to be level, or rising. These are the differences that are expressed in the synthetic continuum in Fig. 2. If subjects perceive three categories, then we expect to find a trimodal distribution corresponding to three perceived categorical differences of the synthetic continuum.

The results from all 10 subjects are shown in the histogram in Fig. 4. This histogram pools their f0-differences between the midpoint of the first syllable /ba/ and the midpoint of following accented syllable /'na/. Negative values on the x-axis indicate f0 in /ba/ was lower than in /'na/ (i.e., a rising pitch) while positive values indicate that f0 in /ba/ was higher than in the following /'na/ (i.e., the pitch was falling). The histogram on the left of the pooled subjects' f0 differences shows some evidence of a bimodal, but not of a trimodal, distribution.

The histogram on the right was derived in the following way. For each subject separately, we calculated the average f0-change across all that subject's /ba'na/ tokens. Then we subtracted this speaker-specific average from the f0-change of that same subject's tokens. (So if for any token, the result of this subtraction is zero, then it means that the f0-change for that token from /ba/ to /'na/ was equal to speaker's average f0-change across these syllables). In this way, we could assess which synthetic stimuli numbers (from st_{01} to st_{16}) were associated with a greater than average f0-drop from /ba/ to /'na/ and which with less.

These results are shown pooled across all subjects in the right panel of Fig. 4. As this figure shows, the first four stimuli were all imitated with approximately the same f0-drop that was much *less* than the average f0-change. The figure also shows that the last 4-5 stimuli were imitated with about the same f0-drop that was much *greater* than the average f0-change. Taken together, these results seem to be indicative of a two-way categorical distinction. Thus, there is clear evidence of differences in imitations to stimuli 1-4 in comparison with stimuli 12-16, but very little evidence of differences in imitations *within* stimuli 1-5 or *within* stimuli 12-16. Moreover, there is more or less an S-shaped progression in the height of the bars across stimuli 1-16 that is very reminiscent of the way in which listeners categorise a synthetic continuum into two categories in segmental perception experiments. Thus, it seems that listeners' imitations are consistent with their perception of two, but not three, tonal categories. We would suggest that one of these is likely to be H* in which the f0-drop from the first to the second syllables of /ba'na/ is not present or minimal, as it is in stimuli 1-4; and the other is likely to be some kind of early peak, most likely H+L* in which the f0-drop is large, as it is in stimuli 12-16. But we do not find much evidence in Fig. 4 in favour of a division of the synthetic continuum into three tonal categories.

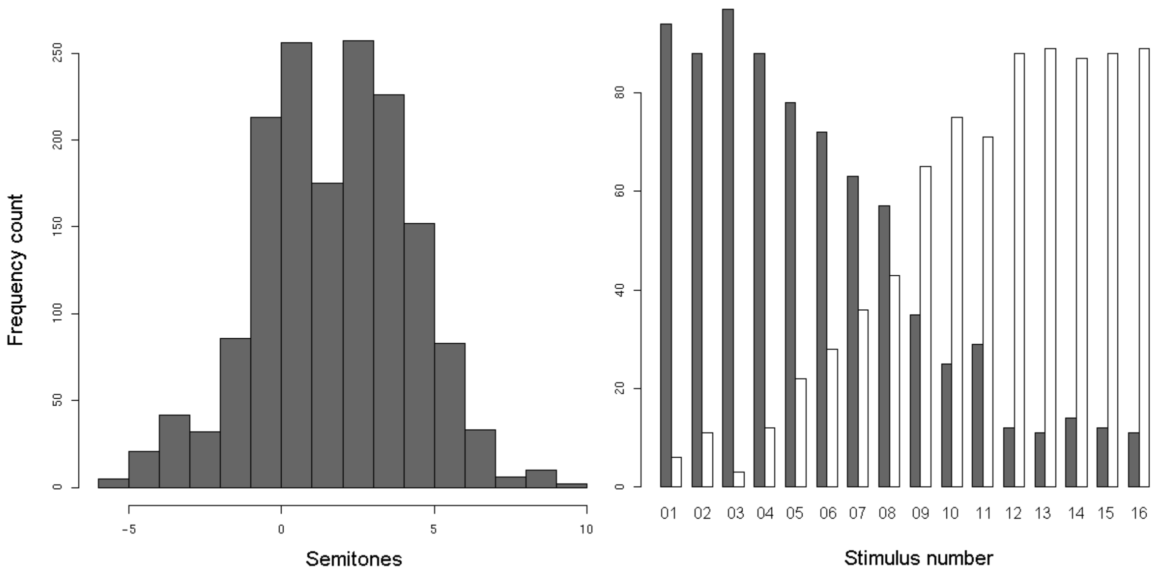

Figure 4: Left: Histogram of P – N where P is the f0 in semitones at the temporal midpoint of initial, prenuclear /a/ in [ba] and N is the f0 in semitones of nuclear /a/ in [na] of 'Bananen' pooled across all speakers. Right: Histogram for all 10 speakers together of $P_k - N_k - m_k$ where P_k and N_k are the same as P and N as defined above but for speaker *k*, and m_k is the mean of $P_k - N_k$ (k= 1, 2, …10 speakers).

3. Discussion and conclusions

Since the results of the perception and production tasks were mostly consistent with a two-category distinction between a mid and some form of early peak, we sought evidence from a large corpus of standard German (the Kiel Corpus of Read Speech, IPDS 1994) for whether phonetic factors might contribute to the perception (and annotation) of f0-contours with H+!H* as opposed to H+L*. The results of this investigation, so far for two speakers, shows firstly that when some form of early pitch-accent is in phrase-final position there is a steep and low f0-fall that is probably accentuated by the L- phrase-tone; and secondly that when early pitch-accents are non-phrase final, i.e., when there are syllables following the nuclear accent in the tail, then the fall across the nuclear accented syllable tends not to be quite so steep or low. Compatibly, there are no examples in the GTOBI training materials where H+!H* is in an intonationally phrase-final position and only one instance where it is final in an intermediate phrase. (Conversely H+L* occurs very often phrase-finally). Thus we would suggest that the extent and rate of the fall of early pitch-accents is phonetically conditioned by their position in the phrase and that it is this phonetic factor that has led some to postulate two distinct pitch-accents H+!H* and H+L*. In Figs. 5 and 6, these potential phonetic differences are illustrated by the different f0-levels in the nuclear accented syllables. For example, in the left panel of Fig. 5, there is an early peak on nuclear-accented *'auch'*, which would be H+L* in GToBI, it is phrase-final, and f0 falls to a low level in the speaker's range. In the right panel, although *'Weißbrot'* is nuclear accented (on *'Weiß'*) and also evidently produced with an early pitch-accent, f0 does not fall to the base of the speaker's range as it does for *'auch'*. So we would argue that these two early peaks are phonetic variants of the same phonological pitch-accent H+L*: that is, the early pitch-accent is realised phonetically as H+L* in *'auch'* but as H+!H* in *'Weiß'* that is not phrase-final and followed by another syllable. The same kind of relationships holds in comparing *'Berlin'* with *'ankomme'* in Fig. 6 (the underlining indicates the rhythmically strongest syllable). Both are again phonologically early pitch-accents, but the fall is steeper and to a lower value in *'Berlin'* in which the early pitch-accent occurs in phrase-final position.

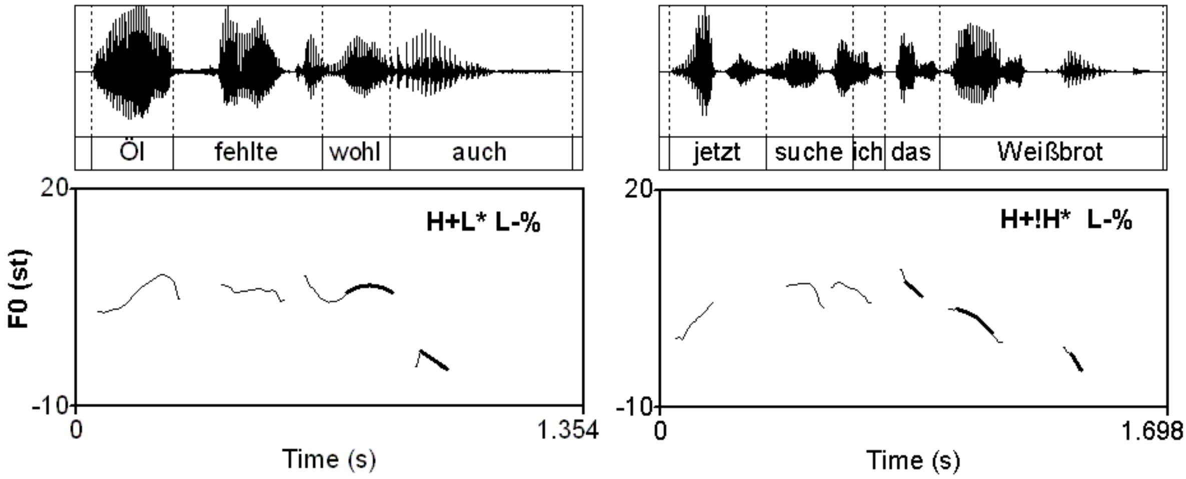

Figure 5: Synchronized time-waveform and f0 trajectories showing tone labels of two read German utterances produced by a male speaker of standard North German (left) of *"Öl fehlte wohl auch."* ('oil was missing too') and (right) *"Jetzt suche ich das Weißbrot."* ('now I'm looking for white bread'). Tonal labels are in accordance with the current guidelines of GToBI (Grice and Baumann 2002).

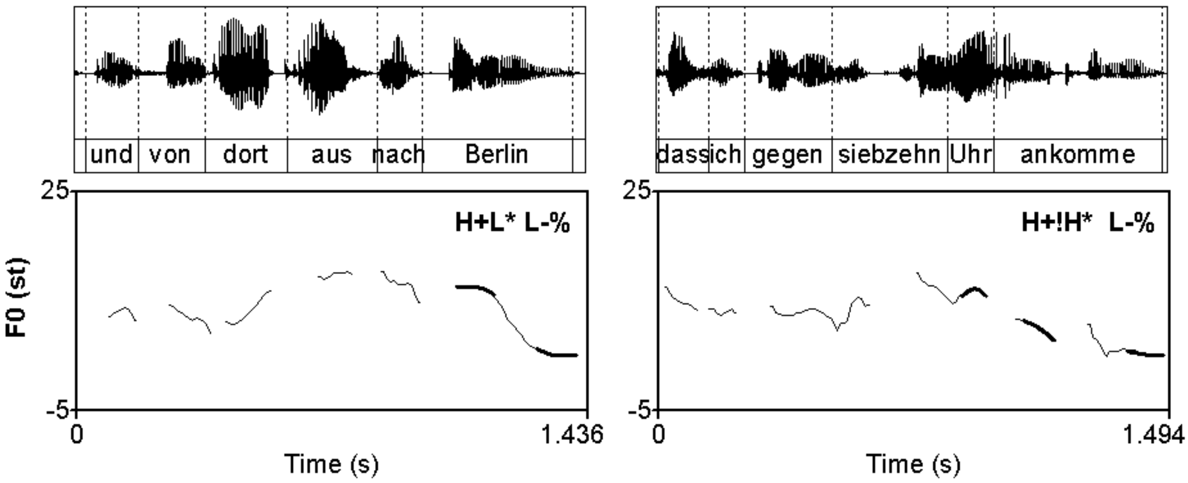

Figure 6: Synchronized time-waveform and f0 trajectories showing tone labels of two read German utterances produced by a male speaker of standard North German (left) of *"…und von dort aus nach Berlin"* ('from there to Berlin') and (right) *"…dass ich gegen siebzehn Uhr ankomme"* ('such that I'll arrive at about 17h'). Tonal labels are in accordance with the current guidelines of GToBI (Grice and Baumann 2002).

Two possible mechanisms might contribute to the lower f0 of early pitch-accents when they are (intonationally) phrase final. Firstly, we can follow Pierrehumbert (1980) and hypothesise that there is an abstract base line which control f0 realisation of L(ow) tones in a phrase, so that the low starred tone on the non-final accented syllables is not realised as low as possible in the speaker's range. Secondly, phrase-final lowering' (Liberman and Pierrehumbert, 1984) is likely to cause f0 to be even lower than would be predicted by declination when the nuclear accented syllable is phrase-final.

In the light of these results, it seems difficult to maintain the idea of a phonological contrast between H+!H* and H+L* in Standard North German while the evidence from the corpus analysis points to a phonetic effect that depends on whether or not the nuclear accent is phrase-final. This is one of the reasons why we would propose not including H+!H* as a phonological tonal category in Standard German. Another is that including bitonal downsteps begs the question of why the inventory should not also be expanded to include bitonal *up*steps like H+^H* or H*+^H. Indeed, if we argue that downstep or upstep are independent phonological entities, then this would be tantamount to expanding the system of contrasts from a two-tone, to a multitonal, paradigmatic system like that of Pike (1945) or Liberman (1975); yet just this type of paradigmatic expansion was heavily criticised (e.g. Bolinger, 1952; Liberman & Pierrehumbert, 1984) in arguing for the primacy of the two-tone autosegmental-metrical model of intonation.

Beyond this issue, the transcription H+!H* implies some kind of downstepped tone, and yet it is being used in a way that is not consistent with the function of downstep in an A-M model. This is because there are quite tight phonotactic restrictions on the occurrence of downstepped tones like !H* and L+!H*: above all, they must always be preceded by a non-downstepped pitch-accent and the only type of pitch-accent that can follow a downstep is another downstepped pitch-accent in the same intermediate phrase.

Whether any of these restrictions are meant to apply to H+!H* is not clear. If not, and then we should abandon the downstepped label in H+!H* and just transcribe it H+M*. However, introducing M* means that the system is founded on a tritonal contrast and this is a radical departure from A-M's two-tone system plus downstep (Gussenhoven, 2004: 104-105).

References

Ambrazaitis, G. 2005 Between fall and fall-rise: substance – function relations in German phrase-final intonation contours. *Phonetica 62,* 196-214.

Beckman, M. and J. B. Pierrehumbert 1986 Intonational structure in Japanese and English. *Phonology Yearbook3,* 255-310.

Benzmüller R., M. Grice, S. Baumann Trainingsmaterialien zur Etikettierung deutscher Intonation mit GToBI. Version 2 (Überarbeitete Fassung). URL1: *http://www.uni-koeln.de/phil-fak/phonetik/gtobi/guidelines-version2c.html*

Bolinger, D. 1951 Intonation: levels versus configurations. *Word 7,* 199-210.

Brosius, F. 2002 SPSS 11. Bonn: mit Verlag.

Dombrowski, E. 2003 Semantic features of accent contours: effects of F0 peak position and F0 time shape. *Proceedings of the XVth International Congress of Phonetic Siences*. Barcelona, 1217-1220.

Féry, C. 1993 German Intonational Patterns. Tübingen: Niemeyer.

Grabe, E. 1998 Comparative Intonational Phonology: English and German. *MPI Series in Psycholinguistics 7.* Wageningen: Ponsen und Looijen.

Grice, M. 1995 Leading tones and downstep in English. *Phonology 12,* 183-233.

Grice, M. and S. Baumann 2002 Deutsche Intonation und GToBI. *Linguistische Berichte 191.* Helmut Buske Verlag, 267-298.

Gussenhoven, C. 2002 Intonation and Interpretation: Phonetics and Phonology. In: *Proceedings Speech Prosody 2002*. Aix-en-Provence. URL2 : *http://www.lpl.univ-aix.fr/sp2002/pdf/gussenhoven.pdf*

Gussenhoven, C. 2004 The Phonology of Tone and Intonation. Cambridge University Press.

Heise, D.R. 1970 The semantic differential and attitude research. In G. F. Summers (ed.) *Attitude Measurement*. Chicago: Rand McNally, 235-253.

IPDS 1994 *The Kiel Corpus of Read Speech.* Volume 1. CD-ROM No. 1. IPDS, Kiel.

Kohler, K. J. 1987 Categorical pitch perception. *Proceedings of the XIth International Congress of Phonetic Siences*. Tallin, 331-333.

Kohler, K. J. 1997 Modelling prosody in spontaneous speech. In Y. Sagisaka, N. Cambell, and N. Higuchi (eds.) *Computing prosody. Computational models for processing spontaneous speech.* N.Y.: Springer, 187-210.

Kohler, K. J. 2005 Timing and communicative functions of pitch contours. *Phonetica 62,* 88-105.

Ladd, D. R. 1983 Phonological features of intonational peaks. *Language 59,* 721-759.

Ladd, D.R. 1996 Intonational Phonology. Cambrigde University Press.

Leonhart, R. 2004 Lehrbuch Statistik: Einstieg und Vertiefung. Göttingen, Hans Huber.

Liberman, M.Y. 1975 The Intonation System of English. Ph.D. dissertation, MIT. [Published by Garland Press, New York].

Liberman, M.Y. and J.B. Pierrehumbert 1984 Intonational invariance under changes in pitch range and length. In M. Aronoff and R.T. Oehrle (eds.) *Language, Sound, Structure: Studies in Phonology Presented to Moris Halle by His Teacher and Students.* MIT Press, Cambridge, 157-223.

Ohala, J. J. 1984 An ethnological perspective on common cross-language utilization of F0 of voice. *Phonetica 41,* 1-16.

Osgood, C.E., Suci, G.J., and P.H. Tannenbaum 1957 The measurement of meaning. University of Illinois Press, Urbana.

Pierrehumbert, J. B. 1980 The Phonology and Phonetics of English Intonation. Ph.D. dissertation, MIT. [Published by Indiana University Linguistics Club, Bloomington].

Pierrehumbert, J. B. and S. A Steele 1989 Categories of tonal alignment in English. *Phonetica 46,* 181-196.

Rathcke, T. and J. Harrington in press Is there a distinction between H+!H* and H+L* in standard German? Evidence from an acoustic and auditory analysis. *Proceedings Speech Prosody 2006,* Dresden.

Silverman, and J.B. Pierrehumbert 1990 The timing of prenuclear high accents in English. *Laboratory Phonology 1,* 72-106.

Uldall, E.T. 1964 Dimensions of meaning in intonation. In Abercrombie, Fry, MacCarthy, Scott, Trim (eds.) *In honour of Daniel Jones.* Longman, London, 271-279.