Prenuclear Rises in Northern and Southern German

Doris Mücke, Martine Grice, Anne Hermes, Johannes Becker

I*f*L Phonetik
University of Cologne, Germany
{doris.muecke; martine.grice; anne.hermes; becker.johannes}@uni-koeln

Abstract

We investigated the effect of dialectal background on tonal alignment in LH rising pitch accents in two varieties of German: those spoken in Düsseldorf and Vienna. We confirm the findings of Atterer & Ladd (2004, [1]) that L and H tones in prenuclear rises are aligned later in Southern than in Northern German varieties, and, more specifically, that both L and H tones in Viennese rises were considerably later than L and H in any of the other varieties. Mean alignment values for both tones in these two varieties can be located on a continuum comprising the values for all four varieties.

1. Introduction

The experiment takes as a starting point the study by Atterer & Ladd (2004, [1]), in which a difference in the timing of prenuclear rising pitch accents was found when comparing two groups of speakers, one from the South of Germany (mainly Bavaria) and one from the North. Although alignment was later for speakers from Southern Germany, the differences were significant only for the beginning of the rise, the L turning point. Atterer and Ladd point out that there was a great degree of inter-speaker variation in the alignment of H, since the groups were fairly heterogeneous, especially the 'Northern German' group.

In an acoustic and articulatory study (Mücke et al., under revision, [2]) we found a similar difference between Southern and Northern varieties, but for another Southern variety (Vienna) and a more closely controlled Northern one (Düsseldorf). However, although we looked at prenuclear accents, our speech material differed from Atterer & Ladd's in one important aspect: the accents were contrastive (since they were on contrastive themes), while Atterer & Ladd's data involved neutral, non-contrastive accents.

The present study set out to replicate the Atterer & Ladd study, investigating neutral prenuclear accents, but in the two varieties already analysed in Mücke et al., [2], Vienna and Düsseldorf. Our motivation was to find out whether the differences we found in contrastive accents across these two varieties are also found when comparing non-contrastive ones, allowing us to shed further light on the question as to whether regional variation of the type found in Atterer and Ladd's study is of a continuous or discrete nature.

2. Method

2.1. Speakers

We recorded eleven native German speakers. For Vienna there were three female and two male speakers, for Düsseldorf there were two male and four female speakers. All speakers grew up in Vienna or Düsseldorf respectively. They were all students in their mid-twenties with an average age of 24 years.

2.2. Speech Materials

Speech materials were based on the Atterer and Ladd sentences, which were designed to elicit bitonal prenuclear pitch accents in non-contrastive contexts. The main stressed syllable of the first content word was the test syllable, and was expected to carry the rise. Every test syllable was surrounded by at least one or two weak syllables to its left and right to avoid stress clash. Every test word was either an adjective followed by a noun or a noun followed by a genitive construction to ensure that the prenuclear rise was on the first content word. An example sentence used in both Atterer & Ladd's and our experiment is given in (1).

(1) Die Er<u>nen</u>nung Meiers zum Minister wurde nicht von allen Parteimitgliedern begrüßt.
The nomination of Meier as minister was not welcomed by all party members (the target syllable is underlined)

Thirteen sentences had test words in prenuclear position. Additionally, ten sentences were constructed with test words in nuclear position. However, we are not able to present results on the alignment of L and H in nuclear accents, since too few utterances carried a rising nuclear pitch accent, e.g. only 18 in the Viennese recordings.

2.3. Recordings and labeling procedures

The Viennese speakers were recorded in Vienna at their homes, while the Düsseldorf speakers were recorded in Düsseldorf at the Heinrich-Heine-University in a quiet room. All recordings were made with a portable DAT-recorder and a condenser microphone. The speakers were asked to read the test sentences from cards in two different pseudo-randomised orders (23 target sentences and 17 fillers). No further instructions were given to the informants. We recorded a total of 880 utterances, including 286 tokens with target prenuclear accents (11 speakers x 13 stimuli x 2 repetitions), which are analysed here.

All recordings were digitised at 44,1kHz/16bit. F0 contours and acoustic waveforms were annotated by hand in EMU. Landmarks were identified along the lines of Atterer & Ladd 2004, [1].

F0 landmarks: F0 values were extracted with a 7.5ms correlation window and a 3ms frame spacing and displayed for hand labelling. Local turning points were identified around the area of the rise contour: the F0 minimum and maximum at the start and the end of the rise. If no local maximum could be identified in the contour for H (which was rare), a clear change of the slope from a steep rise to a plateau or shoulder was used as the point in time the rise starts or ends. If no turning point was identifiable for L (which was often the case for target syllables with only one weak syllable to the left), the utterance was removed from the analysis (since in those cases L placement would have been rather impressionistic in our corpus). The following F0 labels were identified:

L: Low valley at the beginning of the rise.
H: High peak at the end of the rise.

*Segmental landmarks: S*egmental boundaries were identified in the acoustic waveform. For annotation, an oscillogram and wide-band spectrogram were displayed simultaneously. Segmental boundaries for the combinations of nasals and vowels were identified at the abrupt change in the spectra of the nasals at the time the oral closure was formed or released. Abrupt changes in spectra were also observable for laterals and vowels, especially for the intensity of higher formants in laterals. The following boundaries for segmental landmarks were identified:

- **C1:** Start of the onset consonant in the accented syllable.
- **V1:** Start of the vowel in the accented syllable
- **C2:** Start of the consonant following the vowel in the accented syllable.
- **V2:** Start of the vowel in the syllable following the accented one.
- **C3:** End of the vowel in the syllable following the accented one.

An example for segmental and F0 landmarks is given in figure 1 for the test word *Ver<u>läng</u>erung* (extension) in the German sentence *"Die Ver<u>läng</u>erung der Ausleihfrist ist leider nicht möglich"* (An extension of the return date is unfortunately not possible).

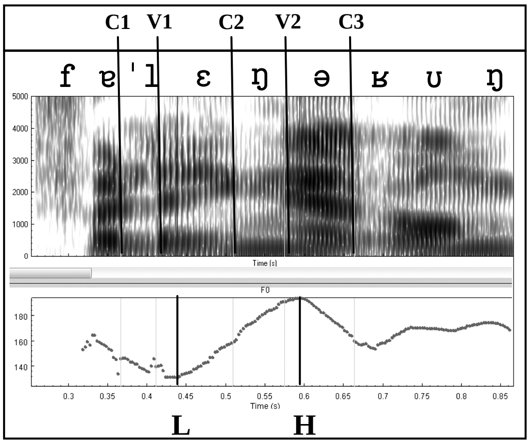

Figure1: *Spectrogram and F0 contour with labels for the test word "Verlängerung", speaker MS (Vienna).*

3. Results

Stimuli: A total of 180 target prenuclear accents were included into the statistical analysis, 97 tokens for Vienna and 83 tokens for Düsseldorf. We removed all stimuli with either mispronunciations, hesitations, utterances carrying a syllabic nasal in the syllable following the test syllable, such as in [nɛnn̩zveɐtɐ] (*<u>nenn</u>enswerter*: noteworthy), or no clear F0 turning point for L in the F0 contour (since in 7 out of 13 test sentences the initial target syllable was preceded by only one (weak) syllable (e.g. *die*), L placement was rather difficult, see 2.2. We excluded a total of 106 tokens (which was nearly one third of the prenuclear corpus).

Analysis: We measured mean values and standard deviations for the following alignment variables.

L-C1: L relative to the beginning of the initial consonant
L-V1: L relative to the beginning of the tonic vowel
H-V2: H relative to the beginning of the posttonic vowel

All mean alignment data for L and H is given for each speaker separately in table 1 (with standard deviation in parentheses). Negative values indicate that the landmark in the F0 contour occurs before the segmental label.

All alignment landmarks were chosen on the basis of the segmental anchor hypothesis in which L and H (measured as local turning points in the F0 contour) co-occur with nearby landmarks in the segmental string (segmental boundaries). We analysed the measures with one-way ANOVAs on the dependent alignment variables (L-C1 or L-V1 or H-V2). Table 1 shows the mean alignment data for L and H relative to the segmental landmarks.

Vienna	L – C1 (ms)	L – V1 (ms)	H – V2 (ms)
1 MS	78 (55)	-4 (59)	33 (24)
2 SR	69 (40)	-8 (38)	39 (38)
3 AV	88 (37)	2 (42)	45 (36)
4 AK	109 (47)	18 (47)	41 (62)
5 NF	141 (41)	49 (34)	94 (40)
Grand mean	97 (50)	11 (48)	50 (45)
Düsseldorf			
1 AW	33 (32)	-35 (28)	6 (20)
2 DOM	36 (27)	-30 (21)	16 (17)
3 SH	35 (29)	-43 (26)	15 (22)
4 AP	28 (14)	-53 (15)	16 (16)
5 SAN	37 (30)	-28 (31)	18 (13)
6 COR	10 (18)	-49 (11)	5 (15)
Grand mean	30 (26)	-40 (24)	13 (17)

Table 1: *Mean latencies for L and H relative to segmental landmarks (beginning of a consonant or vowel in the test syllable, C1 V1, or in the posttonic syllable, V2).*

Results L alignment: Latencies were calculated for L relative to the onset and nucleus of the accented syllable, C1onset and V1onset. While L occurs in the onset consonant of the target syllable (C1) in the Düsseldorf data (on average 30ms after the beginning of C1), it occurs in the vowel of the target syllable (V1) in the Vienna data (on average 11ms after the beginning of V1). We analysed the measures with a one-way ANOVA on the dependent alignment variables L-C1 and L-V1 with DIALECTAL BACKGROUND as an independent variable. DIALECTAL BACKGROUND reached significance for both L alignment variables, L-C1 ([F (1, 180) = 113,467, p<0.001] and L-V1 ([F (1, 180) = 69.303, p<0.001]. The rise starts significantly later in the Viennese data than in the Düsseldorf data. This confirms the findings of Atterer & Ladd that rises in prenuclear pitch accents start later in Southern than in Northern ones.

Results H alignment: H occurred within the nucleus of the posttonic syllable, V2, in both dialects. Latencies were calculated relative to the beginning of V2. For the Düsseldorf speakers, H occurred just after the segmental boundary between the intervocalic consonant and the following vowel (on average 13ms after the boundary C2/V2) and for the Viennese speakers, H was on average 50ms after this segmental boundary. A one-way ANOVA with the independent variable DIALECTAL BACKGROUND and the dependent variable H-V2 reached significance ([$F (1, 180) = 45.412$, $p<0.001$]. Our results then not only confirm the trend for H peaks to be later in Southern varieties, but also provide statistical significance for the difference between a Northern and Southern variety.

Proportional alignment measure: As can be seen from the standard deviations in table 1, we found rather high variability in the alignment for L and H relative to the nearby landmarks (L-C1, L-V1, and H-V2). We thus calculated proportional measurements, since it has been suggested that such measurements might account for some of the variability (Silverman & Pierrehumbert 1990, [5]). We used the index recently proposed by Rathcke & Harrington 2007, [6] for German pitch accents in open and closed syllables, calculating latencies for the tonal target (T) as a function of syllable duration ((T-V1)/(V2-C1)). The syllable duration was calculated for the CVC-string (closed syllable with an ambisyllabic unit). Results are provided in the boxplots in figure 2:

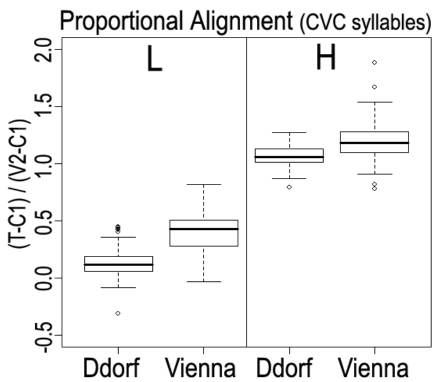

Figure 2*: Proportional alignment of L (left boxes) and H (right boxes) for Vienna and Düsseldorf variety. (T = tone), all speakers.*

For L relative to C1, there is a clear difference observable in the distribution of the boxes. L was later in the Southern variety than in the Northern ones. A one-way ANOVA with (L-C1)/(V2-C1) as dependent variable and DIALECTAL BACKGROUND as independent variable reached significance with [$F (1, 180) = 102.294$, $p<0.001$]. The difference for H was less clear: H was only slightly later in the Vienna data. However, in a one-way ANOVA with the dependent alignment variable (H-C1)/(V2-C1) and independent variable DIALECTAL BACKGROUND, DIALECTAL BACKGROUND was a significant factor ([$F (1, 180) = 39.515$, $p<0.001$]).

L and H compared: In figure 3 we provide a schematic diagram of alignment using mean absolute durations (for five speakers from Vienna and six speakers from Düsseldorf). The figure is to scale and calculated on the basis of mean durations of the segments (C1 V1 C2 V2). As pointed out above, while in the Northern variety, L was aligned with the initial consonant, in the Southern variety it co-occurred with the following vowel in the tonic syllable. Thus, L was aligned with a different segmental anchor, C1, for the North and V1 for the South. However, H co-occurred with the same segment in both varieties: the unstressed vowel V2, although H occurred early in V2 in the Northern variety.

Figure 3: *Alignment of tonal targets with the segmental string in prenuclear accents, all speakers.*

Segmental durations: Atterer & Ladd suggested that the differences in alignment across varieties might have been greatly affected by differences in the duration of the individual segments. However, as can be seen in figure 3 (above) the proportional differences in segmental durations were relatively small, making it unlikely that differences in alignment could be simply due to segmental durations.

	Ddorf	Vienna	p
$C1_{dur\ (ms)}$	69 (19)	85 (17)	***
$V1_{dur\ (ms)}$	68 (18)	86 (30)	***
$C2_{dur\ (ms)}$	59 (13)	65 (27)	n.s.
$V2_{dur\ (ms)}$	55 (21)	65 (13)	***
$V1_{dur}/C1_{dur}$	1.08 (0.48)	1.05 (0.42)	n.s.

Table 2: *Segmental durations in ms and the ratio of the accented vowel to the preceding consonant for Düsseldorf and Vienna, all speakers pooled (for p-values n.s. = p>0.05).*

Table 2 show mean segmental durations and standard deviations (in parentheses) for the segments in the C1V1C2V2 sequences for each variety separately (all speakers together). While Atterer & Ladd found no difference in the mean overall duration of the accented syllable C1V1C2, we found the mean overall duration for the accented syllable C1V1C2 to be 40ms *longer* in Vienna (on average 236ms, 26sd) than in Düsseldorf (on average 196ms, 30sd). The difference reached significance ([$F (1, 180) = 89.142$, $p<0.001$]. In addition, we calculated the ratio of the accented vowel and the preceding consonant (V1dur/C1dur). While Atterer & Ladd found a difference between the varieties in the vowel-consonant ratio (in Southern German, the vowel was longer and the preceding consonant shorter), we found no systematic difference ([$F (1, 180) = 0.145$, $p=0.704$]. These results support our assumption that segment durations are unlikely to be responsible for alignment differences. Contrary to Atterer & Ladd, we found an increase in duration for C1 (16ms longer in Vienna) and V1 (18ms longer in Vienna), as well as an increase in the duration of the unstressed vowel in

Vienna (10ms longer in Vienna). The difference in segmental durations reached significance across the varieties for C1 ([F (1, 180) = 35.794, p<0.001], V1 ([F (1, 180) = 21.432, p<0.001], and V2 ([F (1, 180) = 15.916, p<0.001]. While we found no difference in the ratio of the segments in the accented syllable, we found longer durations for most of the segments produced by the Viennese speakers which might be due to an overall slower articulation rate.

4. Discussion

In contrast to Atterer & Ladd's study, this experiment deals with two homogeneous speaker groups (Vienna and Düsseldorf). Atterer & Ladd found the start of the rise, L, to be aligned significantly later for Southern speakers than for Northern speakers. However, they did not find a significant difference for the alignment of the end of the rise, H. In our recordings, we found a significant difference for both, L and H. The difference reached significance, both for nearby landmarks and proportional measurements.

In both the Düsseldorf and the Vienna varieties, the high tone, H, occurred in the nucleus of the posttonic syllable. This means H falls outside the accented syllable with which it is associated phonologically. In the case of Atterer & Ladd's Northern German and our Düsseldorf variety, H occurred about 13ms after the beginning of V2, which is close to the syllable boundary, which might suggest an alignment with the syllable edge. However, Atterer & Ladd's Southern German and our Viennese German place the peak well into the unstressed vowel, making such an account implausible for these varieties.

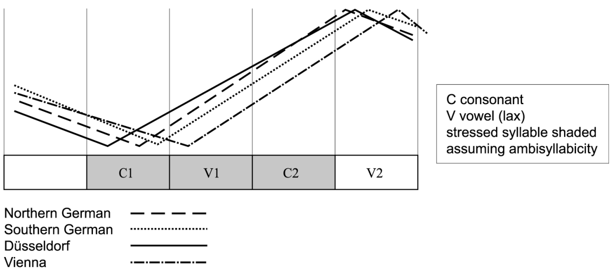

Figure 4: *Summary of alignment properties of prenuclear LH accents. Illustration of German varieties adapted from Atterer & Ladd 2004; Düsseldorf and Vienna added.*

Figure 4 combines our results with the results reported in Atterer & Ladd [1]. The figure is stylised in that all segments were given the same space representing their duration. The figure illustrates the similarity of the alignment pattern between the Düsseldorf variety and Atterer & Ladd's 'Northern German' (L aligned with onset consonant of the accented syllable (C1), H aligned with the beginning of the vowel of the following syllable (V2)). Atterer & Ladd's 'Southern German' and the variety spoken in Vienna both display a later alignment of L and H than Düsseldorf or Northern German. Although both L and H are later in the Vienna variety, the alignment of L is different in a crucial way: whereas in Atterer and Ladd's Southern German L is late – but like the other varieties, still in the onset consonant – in Vienna it is in the vowel (V1), suggesting a qualitative difference in alignment. However, when seen in relation to the other varieties, the later alignment is Vienna appears to lie at the end of a continuum of H alignment points.

Our results are also in line with Braun's recent study (Braun 2007, [3]), comparing Munich in the South and Münster in the North. In her data the sentences were shorter, and the type of nuclear (rhematic) accent had an effect on the alignment such that there was an alignment difference across the two varieties only in contexts where the following accent (which was the rhematic one) was high. Since our test sentences were long enough to contain multiple intermediate phrases and even multiple Intonation Phrases, the rhematic accent did not constitute the following tonal event, thus precluding a direct comparison. The question as to whether there is a category boundary with the valley alignment points (e.g. whether Viennese L which is in the vowel as opposed to the L in the other varieties, which is in the onset consonant) remains open and would have to be addressed in perceptual terms.

5. Conclusion

This experiment found a difference in tonal alignment between two varieties of German, such that the tones in prenuclear rising accents in a Southern variety (Vienna) are consistently aligned later than in a Northern one (Düsseldorf). This was the case not only for the beginning of the rise (L), as already found by Atterer and Ladd when comparing other Northern and Southern varieties, but also for the end of the rise (H). Since our study was a replication of the Atterer and Ladd study, using mostly the same reading materials, we can locate the differences along a continuum of valley and peak alignment including not only our results but also those of Atterer and Ladd.

6. Acknowledgements

The research reported here was funded by the Deutsche Forschungsgemeinschaft (DFG) under the project "The Structure of Tonal Representations Evidence from Tune-Text Synchronisation (STRETTS)", awarded to Martine Grice and Jonathan Harrington. The Viennese recordings and preliminary labelling were carried out by Gudrun Schuchmann for her Masters' Thesis.

7. References

[1] Atterer, M. & Ladd, D.R. (2004). On the phonetics and phonology of "segmental anchoring" of F0: evidence from German. *Journal of Phonetics* 32(2), 177-197.

[2] Mücke, D., Grice, M., Becker, J., Hermes, A. (under revision). Sources of variation in tonal alignment: evidence from acoustic and kinematic data. *Journal of Phonetics*.

[3] Braun, B. (2007). Effects of dialect and context in the realisation of German prenuclear accents. *Proceedings of the 16th ICPhS*, Saarbrücken, Germany, 961-964.

[4] Gussenhoven, C. (2002). Intonation and Interpretation: Phonetics and Phonology, *Proc. Prosody* 2002, 47–57.

[5] Silverman, K. & Pierrehumbert, J. (1990). The timing of prenuclear high accents in English. In Kingston, J. & Beckman, M. (eds.), *Papers in Lab. Phon. I.* Cambridge: C.U.P.

[6] Rathcke, T. & Harrington, J. (2007). The phonetics and phonology of high and low tones in two falling f0-contours in Standard German. *Interspeech* 2007.