Speech Prosody 2008
Campinas, Brazil
May 6-9, 2008

More on the "segmental anchoring" of prenuclear rises: Evidence from East Middle German

Felicitas Kleber and Tamara Rathcke

Institute of Phonetics and Speech Processing
Ludwig-Maximilians-University of Munich, Germany
{kleber; tra}@phonetik.uni-muenchen.de

Abstract

The paper presents results from a production study on the alignment of prenuclear rising accents in East Middle German in which we focus on two research questions: (1) to what extent can an intermediate variety be integrated in the phonetic alignment continuum from south to north as postulated in [2], and (2) to what extent do time pressure factors from the left-hand context influence the stability of tonal alignment. We rearranged the test material used in [2] with respect to unstressed syllables preceding the accented syllable. Our results show that L is aligned earlier in East Middle German than in Northern and Southern German and that left-sided time pressure effects the alignment of L, but not of H.

1. Introduction

In the autosegmental-metrical (AM) theory of intonation, an intonation contour is represented as a sequence of tonal targets: H(igh) tones referring to f0 maxima or *peaks* and L(ow) tones referring to f0 minima or *valleys*. Separately, L and H are monotonal accents, in combination they are bitonal accents which always consist of a starred tone marked by an asterisk and a leading or trailing tone preceding or following the starred tone. A starredness of a given tone represents its phonological *association* to a metrical prominent syllable. Phonetically, association is reflected as the relative timing (or *alignment*) of the corresponding peak or valley and segmental boundaries of the prominent syllable ([12, 15]). The term *segmental anchoring* has been used to describe the phenomenon of stable phonetic alignment to segmental landmarks (such as e.g. the onset of the accented syllable) under varying phonetic conditions. Past research has provided evidence for the relative stability of low tones and a great variability in the alignment of high tones in rising prenuclear or nuclear pitch accents [e.g. [1] for Greek, [2] for German, [7, 14] for Dutch, [13, 19] for English, [17] for Spanish). There is a strong tendency for high tones to be shifted leftwards under time pressure (e.g., proximity to the phrase boundary or another pitch accent). Remarkably, all these studies investigated the phonetic effect of the time pressure from the *right-hand context*. We wonder whether low tones would show less stable alignment (and vice versa high tones more stability) if there is time pressure from the *left-hand context*.

However, the tonal stability has been found language-internally. From the cross-linguistic or even cross-dialectal point of view, it is hard to find stability in the alignment of high and low tones of bitonal accents. Both phonetic ([2]) and phonological ([10]) analyses were carried out to account for these alignments differences between languages or varieties. In a study on Northern (NG) and Southern German (SG) by [2], stable alignment patterns of the prenuclear rise were found within each variety as well as significant differences between the two varieties. The main findings are shown in Fig.1. The two German varieties differ with respect to the alignment of both L and H: in SG, the low tone is aligned significantly later than in NG; the high tone is also aligned later in SG though this difference is not significant. The authors interpreted these results in terms of different phonetic implementations of the same underlying category, the L*+H accent. Furthermore they assumed that phonetic alignment forms a continuum, similar to the VOT continuum [8] with different values in different languages and varieties.

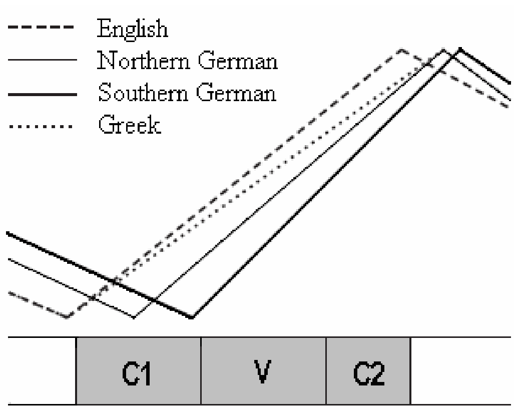

Figure 1: *Schematic representation of the alignment of the prenuclear rise targets in different languages and varieties (after [2]: 187).*

Since the distinction between NG and SG is rather rough, we want to refine the picture by testing a third variety, especially because German dialectologists agree upon a three-way distinction of German dialects on a north-south axis: North, Middle, and South German ([3]). The regional variety chosen for the present investigation is Upper Saxon spoken in East Middle German. Henceforth, we only use the term East Middle German (EG) which is the hypernym for Upper Saxon and which corresponds to the same level of dialectal description as North and Upper German (Southern German in [2]). Both East and West Middle German take up an intermediate position between NG and SG (not only geographically but also linguistically; see for example [3]). Informal auditory observations of EG intonation suggested that prenuclear, but also nuclear accents are always aligned late in the accented syllable. The prenuclear rise sounds like both being very long and reaching the target position quite late (for a detailed description of Saxon intonation see [11]).

The following research questions will be addressed in this paper:

1. Does the alignment of L and H form a phonetic continuum from early to late, in which Upper Saxon takes up an intermediate position between NG and SG because of its dialectal placement linking the two varieties?
2. Is there any effect of the left-hand context comparable to that of the right-hand context? Can we find any evidence for a *rightward* shift of targets depending on the time pressure factor from the context preceding the accented syllable?

2. Method

2.1. Some critic remarks on the method by [2]

The results by [2] are based on absolute segmental duration and alignment values. The syllable durations between NG and SG did not differ drastically in that study, so that the reliability of the data is given. However, a comparison between these results and new data from other dialects is problematic if the durations of the accented syllables are not comparable. As mentioned above, we can assume *a priori* longer syllable durations in EG than in SG or NG. These duration differences would make the comparison of alignment properties in the three varieties impossible. We conclude that it is absolutely necessary to normalize the alignment data with respect to the temporal structure of the segmental string. The normalization method will be explained in 2.5.

2.2. Speech materials and experimental set-up

To test our hypotheses and to allow for a comparison between our and [2]'s data, we used almost the same speech material as in [2]. All test syllables were ambisyllabic and were preceded by one or two unstressed syllables. To test our second hypothesis, we rearranged the test sentences with respect to the number of syllables preceding the accented one. We chose five sentences for two syllable conditions (one/two preceding syllables). Particular attention was paid to the comparable composition of the segments of accented syllables. Three test sentences from the original data by [2] were excluded, because they did not fit into the new experimental requirements. The test material was completed with five test sentences which did not have any syllables preceding the accented one. This procedure resulted in a test corpus consisting of 15 sentences which differed in the number of unaccented syllables (none, one or two) preceding the prenuclear accent. The test material is presented in Table 1. The materials were completed by 19 dummy sentences also taken from [2]. So, the entire experimental set consisted of 34 sentences. They were put in randomized order.

2.3. Speakers and recording procedure

We recorded 12 speakers (7 f, 5 m) aged between 33 and 63. All speakers were born and educated in the East Middle German area and lived most of their lives in that area. All sentences were presented on a laptop computer screen. The subjects were asked to read the sentences at their usual speech rate and with natural accentuation. Subjects were not instructed which intonation/speech melody they should use.

The recording took place in a quiet room at the subjects' homes and was made using the *SpeechRecorder Software* ([9]) and a Sennheisser pc165 USB Headset. The data were digitized at 44.1kHz.

Table 1: *Test words and preceding syllables in the sentences used in the experiment. Accented syllables are printed in bold. Test sentences taken from [2] are indicated by an asterisk.*

	N° of syllables preceding the accented syllable		
	two	one	none
1	*In Er**mang**elung…	Die **mang**elhaften…	**Mang**elhafte …
2	Alle **Minn**esänger…	***Die** **nonn**enhaften…	**Nonn**enhafte …
3	*Die Er**nenn**ung…	*Ein **nenn**ens-werter…	**Nenn**ens-werte…
4	*Die Ver**meng**ung…	*Die **moll**ige…	**Moll**ige…
5	*Die Ver**läng**erung…	*Bei **Läng**engrad…	**Läng**ere…

2.4. Data analysis

The *EMU speech analysis system* ([6]) was used for segmentation of the accented syllable. Tonal labels were set using *praat* ([5]) and then re-converted to EMU. The labeling of tonal targets was done following the core criteria by [2]: The high target (H) was considered to be placed at the midpoint between the local f0 maximum (H2) and the last f0 change in velocity of the rise (H1). The low target (L) was placed at the local f0 minimum of the contour around the onset of the accented syllable. Fig. 2 displays the labeling method in more details. For the measurement of L and H2 we used the automatic f0-detection program implemented in *praat*. H1 labels were set manually. The targeted H was derived automatically in R ([4]). The measurements were done by both authors and checked for consistency.

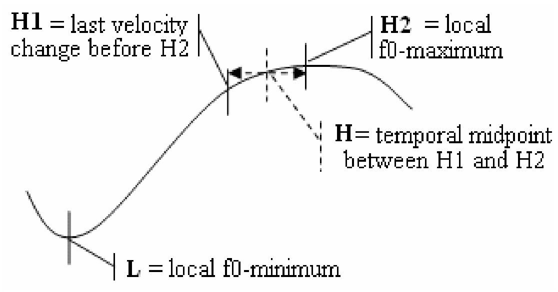

Figure 2: *Schematic representation of measurement points. H1 was measured manually, L and H2 were automatically detected in praat; H was derived automatically in R.*

2.5. Temporal normalization

All statistical analyses were carried out in R. The alignment and segmental data were normalized by means of the equation given in (1). It calculates the linear temporal normalization of the alignment data with relation to the duration of the accented syllable. The syllable onset (S_{on}) is set to 0, the syllable offset (S_{off}) is set to 1. The same normalization method was effectively used in [18] to compare the alignment behavior of high and low tones in two pitch accents:

$$t_T = (T - S_{on}) / (S_{off} - S_{on}) \quad (1)$$

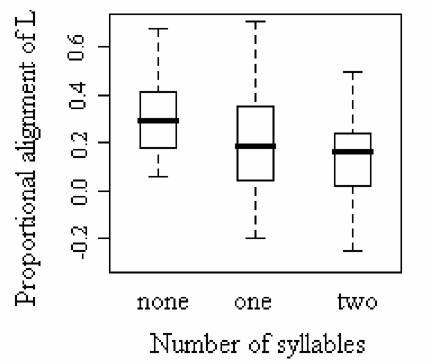

Figure 3: Boxplot for the alignment data of L targets (n_{0s}=n_{1s}=52, n_{2s}=45)

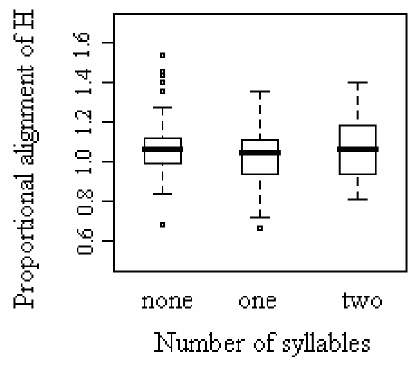

Figure 4: Boxplot for the alignment data of H (n_{0s}=n_{1s}=52, n_{2s}=45)

3. Results

The sentence "Alle Minnesänger…" had to be excluded from the analysis, since most speakers did not accent the test word Minnesänger. Also, the data obtained from the male speaker UEI had to be excluded from the analysis, since his creaky phonation made the necessary f0 measurements impossible.

3.1. Alignment of L and H in EG under left-hand time pressure

Fig. 3 and 4 present the results for proportional alignment of L and H. There is an obvious effect on L-alignment showing that decreasing number of preceding syllables pushed the low target *rightwards*. The effect was absent in the H-alignment data. Two repeated measures ANOVAs (for L and H data separately) were run to test the effect of the independent variable *number of preceding syllables* (none vs. one vs. two) on the dependent variable *alignment.* The analyses were significant for the L alignment ($F_{1,10}$ = 15,0; $p<0.001$) but not for the H alignment. A *post hoc* pairwise t-test showed significant differences in the L alignment between *none-syllable* tokens on the one hand and *one-syllable* ($p < 0.05$) as well as *two-syllable* tokens on the other hand ($p < 0.001$). The difference between *one-syllable* and *two-syllable* tokens was not significant.

3.2. Alignment differences in German varieties

In order to compare the results obtained for EG with those for NG and SG, we normalized the alignment data reported in [2:186] using the same method for the EG data (see 3.5). Table 3 displays actual segmental durations of the accented syllable in the three varieties. The segmental durations in EG are markedly greater than those in SG and NG. Only mean duration data pooled across for three varieties were normalized which was useful for the creation of Fig. 5.

Fig. 5 shows normalized alignment patterns of SG, NG and EG. For EG, only the results of *two-syllable* and *one-syllable* test conditions (see Table 1) were considered. [2]'s results for H as well as L were averaged in order to enable the comparison. The alignment data from the *none-syllable* condition were excluded since there was no comparable sentence structures in [2]'s corpus.

Table 3: *Segmental durations (ms) for three varieties of German.*

Variety	Duration			
	C1	V	C2	Syll
East Middle German	85	92	65	242
Northern German	78	74	56	208
Southern German	71	79	55	205
Mean of all varieties	**78**	**81**	**59**	**218**
Normalized means	**0.36**	**0.37**	**0.26**	**1.0**

As can be seen in Fig. 5, L is aligned earlier in EG than in NG or even SG. There does not seem to be any difference in the H alignment between EG and NG. In both varieties H is aligned earlier, shortly after C2, whereas H in SG is aligned later, i.e. in the segment following C2.

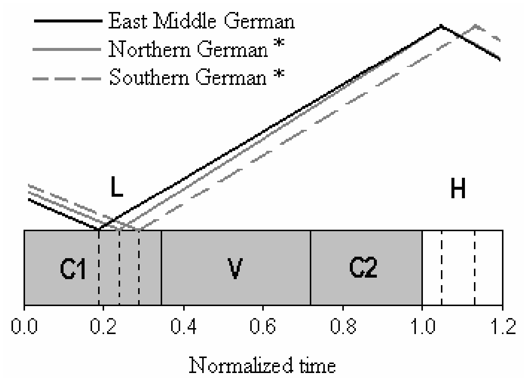

Figure 5: *Schematic representation of the L and H alignment. Data of varieties indicated by * are taken from [2].*

Table 3: *Mean and standard deviations of the proportional alignment of H1, H, and H2 relative to S_{on}.*

SYLL	H1		H		H2	
	mean	sd	mean	sd	mean	sd
none	0.97	0.13	1.08	0.16	1.19	0.21
one	0.91	0.12	1.02	0.14	1.13	0.17
two	0.91	0.12	1.08	0.16	1.25	0.24

3.3. On the stability of H targets

Finally, we turn to the matter of the relative stability of the H alignment found in the EG data. For this purpose, means and standard deviations of the H1, H2 and H alignments were compiled in Table 3. Obviously, H1 showed the same effect of the preceding syllable context as did L (see 3.1.), while H2 was more stable against the time pressure from the left-hand context (which is naturally reflected in the alignment of H).

4. Discussion and Conclusions

With respect to the first research question, we did not find any clear empirical support for the idea of an alignment continuum of both L and H targets in dialect-specific realizations of a prenuclear rise. Instead, the EG data provide evidence for a continual behavior of L, which H does not show. Whereas L is aligned earlier than in other varieties and is placed around the temporal midpoint of the onset consonant, H is aligned shortly after the accented syllable, similar to the H alignment found for NG. To conclude, we found no evidence, that the H of prenuclear rises is aligned later in EG than in the other two varieties. The auditory impression mentioned above (see 1.) can probably be explained by the extremely long (and slow) rise, which emerges from the earlier L alignment and the longer syllable durations as opposed to the other varieties. The alignment difference of H in SG and NG is, according to [2], subtle and not significant. Contrary to the expectation that EG takes up an intermediate position between SG and NG varieties, we found that it is more likely that NG fills this intermediate position. In conclusion, EG extends the alignment continuum for the L tone, but does not lead to the extension of the H continuum. So, the within-linguistic stability of L is in contrast to the between-linguistic variability of L. The opposite is true for the alignment of H.

If we compare the German alignment data for all three varieties to the alignment data of Greek and English, which is schematically shown in Fig. 1, we can at least state that in all German varieties L is aligned within the accented syllable, while in English and Greek it is aligned before the accented syllable. That is, despite the unexpected early L alignment in EG, it has still more in common with other German varieties than with other languages.

With respect to the second research question, our results provide clear evidence that left-sided time pressure factors can influence the phonetic alignment of L: the more time pressure from the left-hand context, the later the alignment. Comparing previous results which showed a relative instability of H tones opposed to stable L tones (e.g. [1, 2, 13, 14, 17]) with present results which show stable H and instable L, one may hypothesize that the temporal proximity of a given tone (H as well as L) to the "source" of time pressure is a relevant predictor for alignment changes. This idea is supported by the data presented for H1 and H2 measurement points.

Our final remark addresses the method: as recognized in some earlier studies ([18, 19]) there is a need for normalization of alignment measurements when comparing data from different languages or varieties. Such comparisons are intended by most AM-studies (e.g. [1, 2, 13, 14]). We think that only normalized data provides a reliable basis for the analysis of alignment patterns under different research conditions and in different languages.

5. References

[1] Arvaniti, A.; Ladd, D.R.; Mennen, I., 1998. Stability of tonal alignment: The case of Greek prenuclear accents. *Journal of Phonetics* 26, 3-25.

[2] Atterer, M.; Ladd, D.R., 2004. On the phonetics and phonology of "segmental anchoring" of *F0*: Evidence from German. *Journal of Phonetics*, 32, pp. 177-197.

[3] Barbour, S.; Stevenson, P., 1990. *Variation in German. A critical approach to German sociolinguistics.* Cambridge: Cambridge University Press.

[4] Bates, D.; Chambers, J.; Dalgaard, P.; Gentleman, R.; Hornik, K.; Iacus, S.; Ihaka, R.; Leisch, F.; Lumley, T.; Maechler, M.; Murdoch, D.; Murrell, P.; Plummer, M.; Ripley, B.; Temple Lang, D.; Tierney, L; Urbanek, S. *The R project for statistical computing*. URL: http://www.r-project.org/ .

[5] Boersma, P.; Weenink, D., 2006. Praat: Doing phonetics by computer (Version 4.5.02). http://www.praat.org/

[6] Bombien, L., Cassidy, S., Harrington, J., John, T., Palethorpe, S., 2006. Recent Developments in the Emu Speech Database System. P*roceedings of the Australian Speech Science and Technology Conference*, Auckland, December 2006.

[7] Caspers, J.; van Heuven, V.J., 1993. Effects of time pressure on the phonetic realization of the Dutch accent – lending pitch rise and fall. *Phonetica* 50, 161-171.

[8] Cho, T.; Ladefoged, P. 1999. Variation and universals in VOT: Evidence from18 languages. *Journal of Phonetics* 27(XXX), 207–229.

[9] Draxler, C.; Jänsch, K., 2004. SpeechRecorder a Universal Platform Independent Multi-Channel Audio Recording Software (Version 2.2.5 beta). http://phonetik.uni-muenchen.de/Bas/software/speechrecorder

[10] Grice, M.; Harrington, J., 2003. Tonal Realization in Two Varieties of German. (Unpublished talk given at the *European Science Foundation Workshop: Experimental Approaches to Tone and Intonation*, University of Oxford: April 2-5, 2003).

[11] Kügler, F., 2005. Swabian and Upper Saxon Intonationa Patterns. (Ph.D. thesis., University of Potsdam.)

[12] Ladd, D.R., 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.

[13] Ladd, D.R.; Faulkner, D.; Faulkner, H.; Schepman, A., 1999. Constant "segmental anchoring" of *F0* movements under changes in speech rate. *Journal of the Acoustical Society of America* 106(3), 1543-1554.

[14] Ladd, D.R.; Mennen, I.; Schepman, A., 2000. Phonological conditioning of peak alignment in rising pitch accents in Dutch. *Journal of the Acoustical Society of America* 107(5), 2685-2696.

[15] Pierrehumbert, J.B. 1980. *The phonology and phonetics of English intonation.* Ph.D. thesis, MIT.

[17] Prieto, P.; van Santen, J.; Hirschberg, J., 1995. Tonal alignment patterns in Spanish. *Journal of Phonetics* 23, 429-451.

[18] Rathcke, T.; Harrington, J., 2007. The phonetics and phonology of high and low tones in two falling f0-contours in Standard German. *Interspeech 2007, Antwerpen.*

[19] Silverman, K.E.; Pierrehumbert, J.B., 1990. The timing of prenuclear high accents in English. In *Papers in Laboratory Phonology* I, J. Kingston; Meckman, M. (eds.). Cambridge: Cambridge University Press, 72-106.