# EFFECTS OF PHRASAL POSITION AND METRICAL STRUCTURE ON ALIGNMENT PATTERNS OF NUCLEAR PITCH ACCENTS IN GERMAN: ACOUSTICS AND ARTICULATION

Henrik Niemann & Doris Mücke

I*f*L Phonetik, University of Cologne, Germany
{henrik.niemann;doris.muecke}@uni-koeln.de

## ABSTRACT

In this production study, we investigate the effects of phrasal position and metrical structure on alignment patterns of nuclear rising pitch accents in German. In the acoustic domain, peak alignment varies across open and closed syllables. In the articulatory domain, these effects of metrical structure disappear. The peak shows a stable coordination pattern with the vocalic target of the accented syllable. However, effects of phrasal position occurred in both domains leading to a leftward shift of the peak towards the accented vowel. We conclude that alignment is best understood as coordinative structures between tones and oral constriction gestures instead of co-occurrence of nearby-landmarks on the acoustic surface.

**Keywords**: peak alignment, variability, coordination, gestures

## 1. INTRODUCTION

Research in tonal alignment is concerned with the arrangement of F0 events with the segmental string, e. g. the alignment of pitch accents with acoustic landmarks such as syllables or vowel onsets. This endeavor tentatively led to the Segmental Anchoring Hypotheses (SAH) stating that F0 valleys (L) and (H) peaks of a rising pitch accent are stably aligned - or anchored - with segments (see [5,8,11] for an overview). However, while it has been shown that L continuously aligns shortly before or after the acoustic onset of the accented syllable [1,10], there is evidence that the alignment of H is more variable in that it is more sensitive to the metrical structure and the position in the intonation phrase, i.e. accentual peaks align earlier in open syllables than in closed ones [10,12] and they align earlier in phrase-final position compared to other positions within the intonation phrase [4,13]. Even though articulatory studies have already shown that accentual peaks align more closely with landmarks in the articulatory domain, these metrical and positional effects are still present [6,10]. Thus, instead of evaluating the co-occurrence of tones with events in the segmental string or the articulation, we rather seek out to find stable coordination patterns. In this study, we investigate the alignment of rising nuclear pitch accents relative to both nearby acoustic landmarks and non-nearby landmarks in the articulatory domain.

## 2. METHOD

### 2.1. Articulatory recordings

Four native speakers, all female and aged between 26 and 32 years, participated in this study. All speakers grew up north of the Benrather isogloss and did not report on any hearing or speaking deficits. Articulatory data were sampled with a 3D Electromagnetic Articulograph (Carsten AG 500) at 200 Hz and smoothed with a 3-step-floating mean. The acoustic data were recorded with a time-synchronized microphone and digitized at 16 kHz. Articulatory movements were tracked by sensors on the upper and lower lips as well as on the tongue tip (1 cm behind the tip) and tongue body (two sensors, 3 cm and 4 cm behind the tip). In sum, we recorded 672 tokens (4 speakers x 7 target words x 3 phrasal positions x 8 repetitions).

### 2.2. Speech material

The speech material consisted of seven target words with either /ma:/ or /mam/ as accented syllable. Target words with an open syllable included /ma:/, /’ma:mi/ and /li’ma:ni/. Target word with a closed syllable included /mam/, /’mami/, /’mamzi/ and /’mamila/ (ambisyllabicity is assumed in /’mami/ and /’mamila/). All words were placed in carrier sentences designed to elicit rising nuclear pitch accents in (a) phrase-initial, (b) phrase-noninitial and (c) phrase-final position. The following carrier sentences with the target word /ma:mi/ exemplify our corpus. Squared brackets indicate prosodic boundaries. Target words are bold and underlined.

(a) [Dann dachte sie:]$_{IP}$ [**<u>Mahmi</u>** manipulierte die Bremse.]$_{IP}$
[Then she thought:]$_{IP}$ [Mahmi manipulated the brakes.]$_{IP}$

(b) [Die **<u>Mahmi</u>** manipulierte die Bremse.]$_{IP}$
[The Mahmi manipulated the brakes.]$_{IP}$

(c) [Sie sah dann die **<u>Mahmi</u>**.]$_{IP}$
[Then she saw the Mahmi.]$_{IP}$

All sentences served as answers in small question-and-answer pairs designed to elicit a rising nuclear accent on the target word. These pairs were presented to the subjects on a computer screen. Subjects were prompted to read the answer in a comfortable and natural way after the question was presented auditorily.

### 2.3. Measurements

Articulatory and acoustic data were labelled manually using the EMU speech database system and PRAAT, respectively [3,7]. In the F0 trace, we identified local turning points for L (F0 min) and H (F0 max) in the vicinity of the rising nuclear LH pitch accent. In the acoustics, we annotated segmental boundaries by means of the waveform and oscillogram. In the articulation, we labelled consonantal gestures via the Lip Aperture Index [2]. We identified local minima and maxima in the interlip distance trace corresponding to the opening and closing gesture of the bilabial consonants. The target of the vocalic gesture was identified by means of a local minimum in the vertical tongue body trace corresponding to a zero-crossing in the velocity derivation. The following lags were computed in the acoustic (measure 1-2) and the articulatory domain (measure 3-4, see also figure 4-5 in the discussion).

**1. LtoC1ons:** The beginning of the accentual rise (L) relative to the beginning of the accented syllable.
**2. HtoV1end:** The end of the accentual rise (H) relative to the end of the accented vowel. In open syllables, this segmental landmark denotes the end of the accented syllable. In closed syllables, it is the onset of the coda/ambisyllabic consonant.
**3. L-targC1:** Beginning of the rise relative to the articulatory maximum of the closing gesture in the word-initial consonant.
**4. H-targV:** End of the rise relative to the articulatory target of the vocalic gesture corresponding to the accented syllable.

## 3. RESULTS

### 3.1. Alignment of L

#### *3.1.1. Acoustics*

For the alignment of the beginning of the accentual rise, we measured the F0 min relative to the onset of the accented syllable. An overall rmANOVA with speaker as random factor did not reveal an effect of POSITION [$F(2,6) = 2.31$, $p>0.05$] but of TARGET WORD [$F(6,18) = 3.22$, $p<0.05$] and an interaction [$F(12,36) = 4.25$, $p<0.001$]. This interaction is confirmed by Tukey's post hoc comparisons showing that only in phrase-initial position, L in the target word /li'**ma:**ni/ is significantly aligned later than all other target words. Taken these results together, our data confirm that the beginning of the rise is not affected by the phrasal position nor by the syllable structure and thus exhibits a stable timing pattern relative to the beginning of the accented syllable.

#### *3.1.2. Articulation*

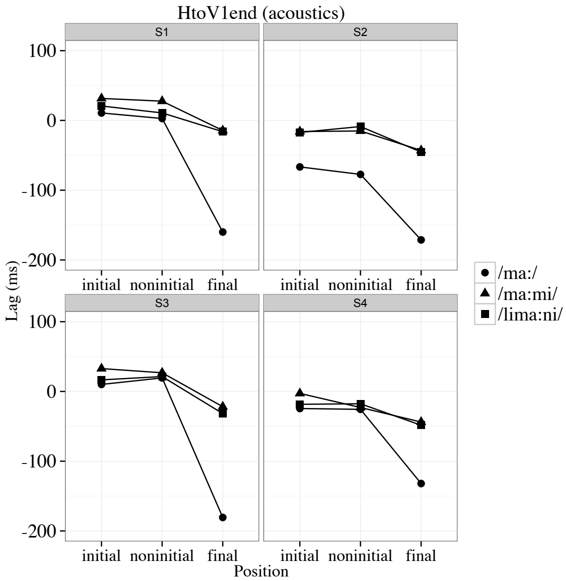
Even though there is trend for two of four speakers to align L earlier in phrase-final condition, the rmANOVA failed to find any effect of TARGET WORD [$F(6,18) = 1.60$, $p>0.05$] or POSITION [$F(2,6) = 4.97$, $p>0.05$]. These results indicate a stable coordination of the beginning of the accentual rise relative to an articulatory landmark, that is the maximum closure for the onset consonant /m/.

### 3.2. Alignment of H

#### *3.2.1. Acoustics*

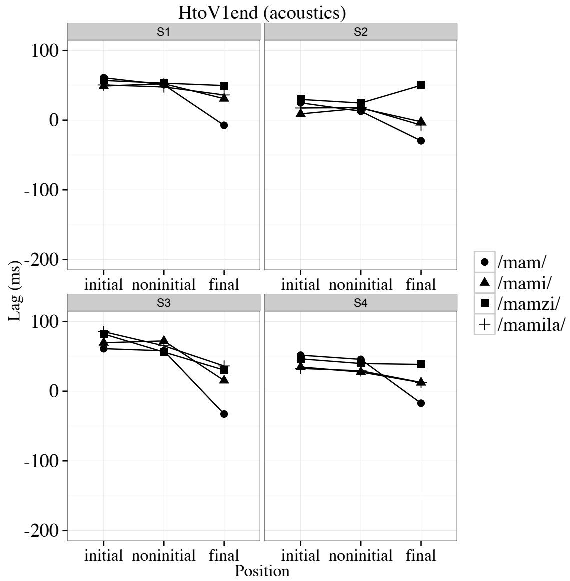
For the alignment of the end of the accentual rise, we measured the F0 peak relative to the end of the accented vowel. As expected, the rmANOVA revealed both an effect of TARGET WORD [$F(6,18) = 75.65$, $p<0.001$] and POSITION [$F(2,6) = 27.49$, $p<0.001$] as well as an interaction [$F(12,36) = 28.06$, $p<0.001$]. Data were thus separated by syllable structure. Figure 1 and figure 2 display means for target words with an open and closed syllable, respectively.

**Figure 1**: Means for the end of the accentual rise to the end of the accented vowel (open syllables).




For both open and closed syllables, Tukey's post hoc comparisons revealed no significant difference between the phrase-initial and phrase-noninitial condition (p>0.05). However, both positions differ from the phrase-final condition in that here the F0 peak aligns earlier (p<0.001). In this position, the rmANOVA revealed an effect of target word [F(2,6) = 71.36, p<0.001]. Tukey's post hoc tests confirmed that in all monosyllabic and all disyllabic target words with an open syllable, the F0 peak aligns earlier in phrase-final condition. As for the disyllabic words with a closed syllable, post-hoc tests showed that in /mamzi/ the peak *was not* shifted leftwards away from the boundary in phrase-final position.

**Figure 2**: Means for the end of the accentual rise to the end of the accented vowel (closed syllables).




To sum up the acoustic results, we found a variable alignment for H in that it aligns earlier in open than in closed syllables. We also found a positional effect in that H was aligned earlier in phrase-final condition. However, this effect varied as a function of syllable structure: It was strongest in monosyllables but weaker in disyllables. Table 1 provides means for the peak relative to the end of the accented vowel.

**Table 1**: Means and standard deviations (in ms) for H relative to the end of the accented vowel.

| HtoV1end | initial | noninitial | final |
|---|---|---|---|
| **/ma:/** | -18 (37) | -20 (42) | -161 (21) |
| **/ma:mi/** | 11 (25) | 4 (27) | -31 (15) |
| **/li'ma:ni/** | 0 (21) | 1 (18) | -36 (14) |
| **/mam/** | 50 (17) | 42 (20) | -22 (12) |
| **/mami/** | 40 (25) | 42 (25) | 14 (14) |
| **/mamzi/** | 54 (22) | 43 (14) | 42 (9) |
| **/mamila/** | 46 (29) | 40 (21) | 20 (21) |

### *3.2.2. Articulation*

Figure 3 and figure 4 display the alignment of H relative to the articulatory vowel target in open and closed syllables, respectively. Note that in open syllables, /ma:/ had to be excluded from this analysis as the articulatory vowel target could not be reliably detected.

**Figure 3**: Means for the end of the accentual rise to the articulatory vowel target /a/ (open syllables).

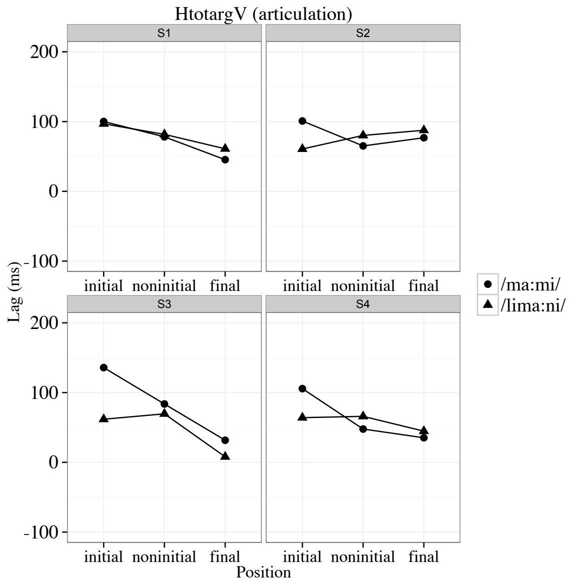
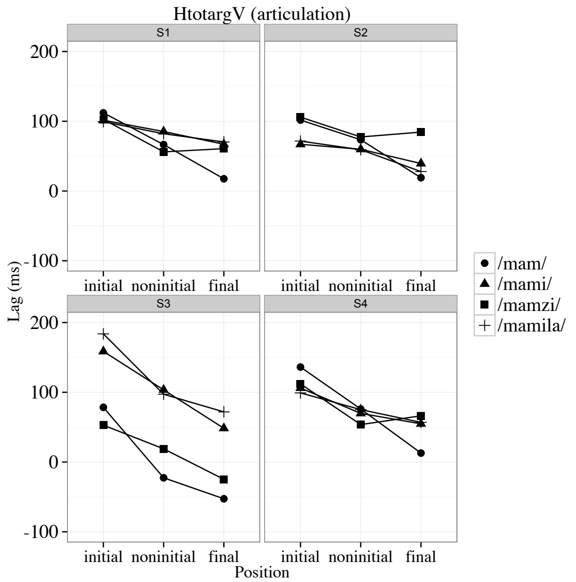



**Figure 4**: Means for the end of the accentual rise to the articulatory vowel target /a/ (closed syllables).




The rmANOVA revealed *no* effect of TARGET WORD [F(5,15)= 0.82, p>0.05], i.e. the peak alignment relative to the articulatory vowel target was neither affected by the syllable structure nor the number of syllables. However, POSITION [F(2,6)=15.85, p<0.01] reached significance as well as an interaction [F(10,30)=6.47, p<0.001]. Pairwise comparisons showed that, except for /lima:ni/, in all target words the peak significantly aligns earlier in phrase-final condition.

To sum up the articulatory results, we have found no significant difference in peak alignment between open and closed syllables or between mono- and disyllabic target words. However, the peak alignment was affected by position in the intonational hierarchy, i.e. the peak was aligned earlier in phrase final position (especially in monosyllabic words). Table 2 provides means for the peak relative to the articulatory vowel target

**Table 2**: Means and standard deviations (in ms) for H relative to the articulatory vowel target /a/.

| HtotargV | initial | noninitial | final |
|---|---|---|---|
| **/ma:mi/** | 111 (17) | 69 (16) | 47 (21) |
| **/li'ma:ni/** | 71 (17) | 74 (7) | 50 (33) |
| **/mam/** | 107 (24) | 48 (47) | -1 (35) |
| **/mami/** | 108 (38) | 80 (19) | 52 (12) |
| **/mamzi/** | 93 (27) | 52 (24) | 47 (49) |
| **/mamila/** | 113 (49) | 78 (16) | 57 (20) |

## 4. SUMMARY

In this paper, we investigated the coordination of rising nuclear pitch accents with acoustic and articulatory landmarks. The beginning of the accentual rise exhibits a stable alignment both relative to acoustic and articulatory landmarks. In contrast, the end the accentual rise is more variable in the acoustic than in the articulatory domain. Acoustically, the peak is affected both by the metrical structure and the phrasal position, i.e. the F0 peak aligns earlier in closed syllables and earlier in phrase-final position. Articulatory, metrical effects *disappear* by relating the peak to the articulatory vowel target. However, a positional effect remains in the phrase-final position as due to an upcoming boundary, the peak is shifted leftwards, especially in monosyllabic words.

Figure 5 illustrates our findings for open and closed syllables (date taken from speaker S4). From top to bottom the figure schematizes the F0 contour, the segmental string of the target word (accented syllables are shaded in grey), the lip aperture and the vertical movement of the tongue body). On the acoustic surface, there is a variable peak alignment in that in /ma:mi/ (open syllable) the peak aligns 3 ms *before* the vowel offset, whereas it aligns 46 ms *after* the vowel offset in /mamzi/ (closed syllable). However, this effect disappears when relating the peak to the articulatory vowel target. Here, it exhibits a stable interval from the articulatory vowel target in /ma:mi/ (106 ms) and /mamzi/ (112 ms).

**Figure 5**: Peak alignment in open and closed syllables (speaker S4).

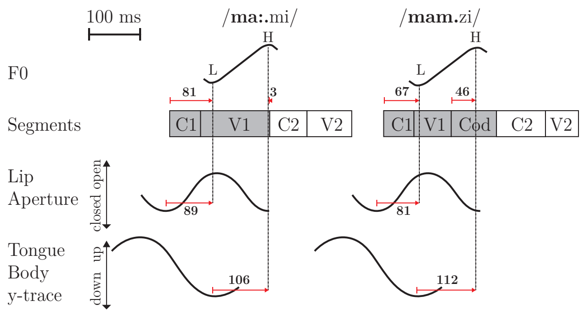
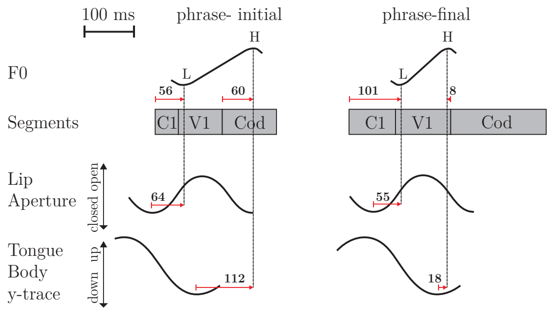



Figure 6 illustrates the positional findings (data taken from speaker S1). In both the acoustic and articulatory domain, the peak is shifted leftwards in phrase-final condition. Acoustically, the peak aligns 60 ms *after* the vowel offset in phrase-initial position, whereas it aligns 8 ms *before* the vowel offset in phrase-final position. This shift is also reflected in the alignment relative to the articulatory vowel target in that the lag between the vowel target and the peak is significantly shrinked from 112 ms in phrase-initial to 18 ms in phrase-final position.

**Figure 6**: Peak alignment in phrase-initial and phrase-final position (speaker S1).




In this study, we have shown that F0 peak alignment in the acoustic domain varies across open and closed syllables. When investigating alignment in terms of nearby segmental landmarks, peaks occur earlier in open syllables than in closed syllables. However, this effects disappears when investigating the coordination between F0 peaks and articulatory gestures. More specifically, the peak shows a stable coordination pattern with the vocalic target of the accented syllable. This supports the findings in [9] reporting a stable coordination pattern of the F0 peak with the transvocalic opening during the production of the accented syllables. However, this coordination pattern (coordination between F0 peak and accented vowel) is affected by an upcoming boundary leading to a leftward shift of the accentual peak towards the vowel target.

## 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] Byrd, D. 2000. Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica, 57 (1),* 3-16.

[3] Boersma, P., Weenink, D. 2013. *Praat: doing phonetics by computer* [Computer program]. Version 5.3.86.

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

[5] D'Imperio, M. 2011. Tonal Alignment. In: Cohn, A.C., Fougeron, C., Huffman, M.K. (eds), The Oxford Handbook of Laboratory Phonology, Oxford: OUP: 265-275.

[6] D'Imperio, M., Espesser R., Lœvenbruck, H., Menezes, C., Nguyen, N., Welby, P. 2007. Are tones aligned to articulatory events? Evidence from Italian and French. In: Cole, J., Hualde, J.I. (eds.). *Laboratory Phonology 9*, 577-607.

[7] Harrington, J. 2010. *The Phonetic analysis of Speech Corpora*. Wiley-Blackwell.

[8] Ladd, D. Robert $^{2}$2008. Intonational Phonology. Cambridge, MA: Cambridge University Press.

[9] Mücke, D., Grice, M., Hermes, A. 2008. The vowel triggers the tone: Evidence from German. Special Session: Coordination of speech gestures: Models and data. 8th Phonetic Conference of China 2008, Beijing, China.

[10] Mücke, D., Grice, M., Becker, J., Hermes, A. 2009. Sources of variation in tonal alignment: evidence from acoustic and kinematic data. *Journal of Phonetics 37 (3),* 321-338.

[11] Prieto, P. 2011. Tonal alignment. In: Van Oostendorp, M., Ewen, C., Hume, B., Rice, K. (eds.), Companion to Phonology, Wiley-Blackwell. 1185-1203.

[12] Prieto, P., Torreira, F. 2007. The Segmental Anchoring Hypothesis revisited: Syllable structure and speech rate effects on peak timing in Spanish. *Journal of Phonetics 35*, 437-500.

[13] Silverman, K.E.A., Pierrehumbert, J. 1990. The timing of prenuclear high accents in English. In: Kingston, J., Beckman, M. (eds), *Papers in Laboratory Phonology I: Between the grammar and Physics of Speech.* 72-106.