Constant Tonal Alignment in Swedish Word Accent II

Svensson Lundmark, Malin

Published in:
The Proceedings of Speech Prosody

2014

Constant Tonal Alignment in Swedish Word Accent II

Malin Svensson Lundmark

Linguistics and Phonetics, Lund University, Lund, Sweden
Malin.svenssonlundmark@gmail.com

Abstract

Studies on accentual tonal alignment of intonation languages suggest that L in rising (LH) pre-nuclear accents anchors with a specific point in the segmental string, while the timing of H varies. This study investigates if lexical accents, too, exhibit a constant alignment by testing the South Swedish word Accent II. When under the strain of tempo variability the L-target was found not to be anchored with syllable onset. The results were not fully conclusive regarding H, but no clear evidence was found against anchoring of H, which could mean that H is an important phonological event in Accent II, while L is not.

Index Terms: tonal alignment, segmental anchoring, word accent, pre-nuclear accent, speech rate

1. Introduction

Over a period of 15 years there has been an on-and-off debate within intonational phonology on whether or not accentual tonal targets (L, H) are constantly aligned with the segmental string. Most studies have focused on pre-nuclear rising accents in intonation languages and have found conclusive results on the start of the rise, the L-target. So far, a language with lexical accents has not been taken into account in the recent research on constant tonal alignment.

1.1. Tonal alignment

Tonal alignment might be seen upon as a wider notion for other concepts such as timing, tonal association or segmental anchoring. The *segmental anchoring principle* presupposes that tonal targets are constantly aligned, and thus anchored at specific points in the segmental string [1]-[3]. Studies that second the principle have focused on rising pre-nuclear accents in intonation languages. Previous studies in the field of constant tonal alignment have displayed an unambiguous case of the tonal target L aligning with syllable onset in pre-nuclear accents, though the precise timing seems to vary across languages. Results include the L-target occurring just before the onset of the accented syllable ([1] for Greek; [4] for Italian), at syllable onset ([5] for Dutch; [3] for English; [2] for German), or after syllable onset ([6] for Mandarin).

While the studies show anchoring of L with the beginning of the syllable, the same consistent result does not exist for the H-target. Some studies found H anchoring after syllable offset [1], [2], [6], or somewhere late in the syllable [3]. Caspers and Van Heuven [5] found that the end of the rise, the H-target, varied considerably under time pressure and thus rejected that it did anchor with the segment. However, they did also consider whether or not the variation had to do with segmental structure.

It can be concluded that the L-target appears to be more stable than the H-target in rising pre-nuclear accents in intonation languages. Niemann et al. [4] has suggested that the cross-linguistic variation of anchoring of L is systematic and can be explained by different phonological structures between the intonation languages. The inconclusive results on H have been highlighted by Niebuhr et al. [7] who addressed the effect of individual speaker strategies. They also proposed that instead other related features are responsible for the consistent results of segmental anchoring.

1.2. Swedish accents

In the prosodic typology of Swedish intonation provided by the Lund Model [8], [9], the two Swedish word accents are assumed to be represented by a fall associated with the stressed syllable in a prosodic word, where the two accents differ in the timing of the fall. There is a regional variation between the Swedish dialects. For example in the South Swedish dialect (South) both accents are timed considerably later than in the Central Swedish dialect (Svea): in South Swedish the high level in Accent I is associated with the stressed vowel and in Accent II with offset of the stressed syllable.

The original dialect typology has later been revised by Bruce [9], who identified, for all dialects, an LHL tonal gesture from which bitonal gestures are extracted; either a fall, H+L, or a rise, L+H. Bruce generalized for Accent II an association of a fall in the dialects with an early timing of the accents (Svea and Göta) and a rise in the dialects with a late timing (South, Gotland, Dala, North). For the South Swedish dialect, a late timed dialect type, Bruce made the specific assumption of a fall, an H+L pattern, for Accent I and a rise, an L+H pattern, for Accent II. The rise in South Swedish Accent II has indeed been shown to be relevant from a perceptual point of view [10].

As if by chance, there is a rough phonetic match between the timing of the lexical Accent II in South Swedish and the pre-nuclear accents in the already mentioned studies on tonal alignment in intonation languages. The research question formulated here is whether or not additional phonetic features are similar such as if L and H are anchored with a segment, as is assumed by the segmental anchoring principle.

The present study is a production study where the hypothesis of segmental anchoring is tested on the Swedish word Accent II. Speech rate is used as an experimental tool and is based on the idea that speakers will try to retain primary features of phonological properties, while they will let other features be modified under time pressure [5]. Speech rate has been used successfully by a number of researchers in studies concerning tonal alignment [3], [5], [6]. Because the Lund Model (revised by Bruce [9]) assumes that both L and H in the rising L+H gesture of Accent II are phonologically relevant, anchoring of L and H is expected.

2. Method

2.1. Speakers and recording

The material was initially collected for a different study in which two age groups were recorded. For this study only the older speakers were tested due to technical issues. There were seven speakers, four males and three females, and the average age of the speakers was 72 years. All speakers were voluntary and spoke the same variety of the South Swedish dialect. A criterion for speaker selection was that they had all lived most of their lives in the same area in the northeastern part of the South Swedish region. Moreover, their parents also had to have lived most of their lives in the area.

All of the recordings were made in people's homes. An IMG Stage boundary microphone (table-microphone) with phantom power was used (ECM-302B) since it is non-invasive and the speakers were expected to be naïve with no prior recording experiences.

The material was read twice by each speaker at three different speech rates: normal, slow and fast. The recording leader set the pace of the speech rate with the leading question and the speaker was asked to answer the question and to follow the speech rate of the recording leader.

2.2. Speech materials and data processing

The materials consisted of three test sentences with the same test word: *många*, meaning 'many' in English. The materials were mixed with 37 further sentences not investigated here. The test word, in its three sentence contexts, fits the following criteria: an unbroken tonal curve, a word Accent II, identical segmental surroundings ([9 syllables] bisyllabic target word [2 syllables]) and that neither syllable and vowel onset, nor syllable and vowel offset coincided. The leading question sought to that the target word *många* [ˈmɔŋːa] 'many' occurred before nuclear accent in all three sentences. The F0 contour of an example sentence can be seen in Figure 1.

The author performed segmentation and annotation in Praat [11]. Since each speaker was recorded twice, the material consisted of 126 items (3 sentences x 2 repetitions x 3 speech rates x 7 speakers). Each target word was segmented into syllables, and in addition, the boundaries of the accented vowel were determined. The boundary between the two syllables was defined as the temporal midpoint of the long, ambisyllabic consonant. The tonal curve has been semi-automatically annotated for the tonal targets L and H by the author (Figure 2). Extracted measures were the start and the end of the rise (L and H), syllable onset and syllable offset.

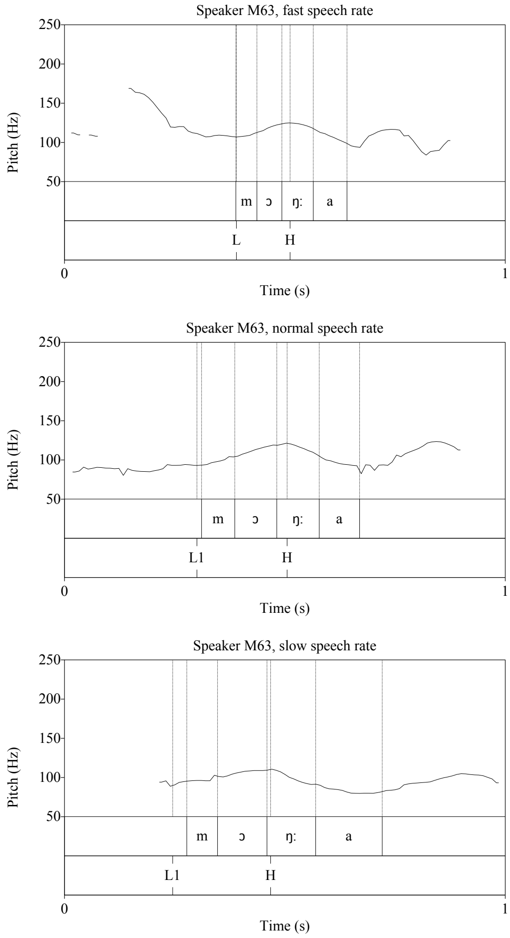

Figure 2: *The segmented target word [ˈmɔŋːa] in three different speech rates, spoken by male speaker M63. The low (L/L1) and the high target (H) of the tonal curve has been annotated.*

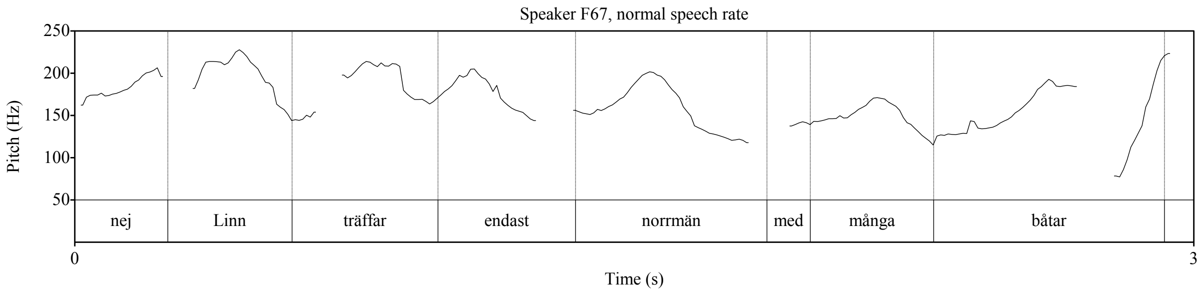

Figure 1: *One of the three target sentences in the material as spoken by female speaker F67 in normal speech rate. The sentence translated into English is: 'No, Linn only meets Norwegians with many boats'.*

3. Results

Average syllable duration shows a difference in speech rate between the recordings (Figure 3). An ANOVA confirmed that speech rate had a significant effect on syllable duration (F = 66.490, df 1, p < .001), concluding that the rate manipulation was successful. Since the segments are affected by speech rate, the temporal distance between the tonal targets L and H should also be affected by speech rate, if they are anchored with the segmental string. An ANOVA was run and showed a significant effect of speech rate on the temporal distance between L and H (henceforth, rise time) (F=17.129, df 1, p = .006) resulting in shorter rise times for faster speech.

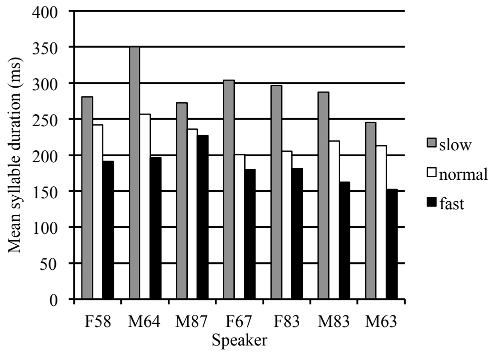

Figure 3: *Average syllable duration for each speaker in each speech rate.*

If anchoring of the tonal targets with specific points in the segmental string occurs this would necessitate a correlation between segment duration and distance between tonal targets. This was tested by means of a Correlations Pearsons (2-tailed) test. There appears to be a weak to moderate positive linear relationship (R= 0.433, N= 74), which indicates a correlation. However, it does not seem to be a convincingly strong correlation (Figure 4).

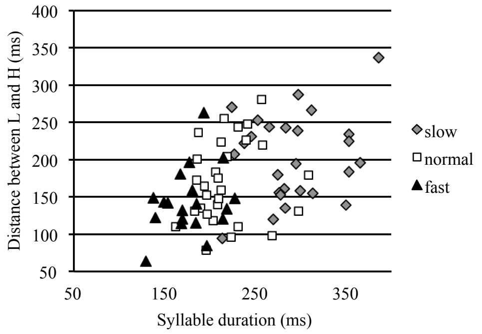

Figure 4: *Scatter plot of the relationship between syllable duration and the distance between tonal targets L and H.*

To test whether the weak to moderate relationship might indicate that either only one or neither of the targets is anchored, two new measures were calculated: the distance between L and syllable onset, and the distance between H and syllable offset, where anchoring is measured as distance in milliseconds. An ANOVA was first run with speaker as random sample. To account for missing data, an average value was first calculated for each speaker across the available items for each condition. The ANOVA showed no significant effect of speech rate on the distance between the tonal target L and syllable onset (F =0.460, df 1, p = .523). An ANOVA was also run on the distance between H and syllable offset, showing no significant effect of speech rate on anchoring of H (F = .702, df 1, p = .434). However, Figure 5 sends a different message and displays a large variance of anchoring of L between the speakers comprising alignments before as well as after syllable onset. The H-target also varies considerably in its alignment, but somewhat less than the L-target (see also Table 1). Notably, the H-target is consistently aligned before syllable offset, on average around 40 ms for both normal and fast speech (Table 1).

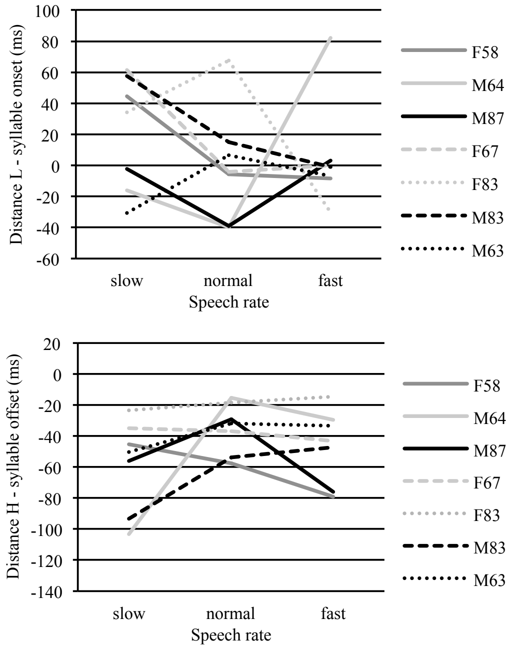

Figure 5: *Distance between tonal target and segment boundary, average value for each speaker. The first graph shows distance between L and syllable onset, the second between H and syllable offset. A negative number indicates that timing is before the boundary.*

Table 1: *Distance between tonal target and segment boundary (ms), average value for each speaker. Standard deviation in parentheses. A negative number indicates that the target is before the boundary. * Only one item in this condition.*

	L - syllable onset			H - syllable offset		
	slow	normal	fast	slow	normal	fast
F58	45 (47)	-6 (97)	-9 (1)	-45 (18)	-58 (14)	-79 (32)
M64	-16 (22)	-40 (*)	82 (*)	-103 (49)	-16 (*)	-29 (*)
M87	-2 (43)	-39 (44)	3 (*)	-56 (32)	-29 (48)	-76 (*)
F67	62 (56)	-4 (9)	0 (4)	-35 (21)	-37 (10)	-43 (17)
F83	34 (*)	67 (39)	-30 (28)	-24 (*)	-19 (13)	-15 (18)
M83	57 (26)	15 (38)	-1 (45)	-93 (9)	-54 (20)	-47 (7)
M63	-31 (36)	7 (29)	-7 (8)	-50 (26)	-32 (25)	-33 (31)
All	24 (51)	9 (54)	-6 (36)	-65 (37)	-38 (23)	-40 (27)

The anomaly of only one available item in some conditions for three of the speakers (M64, M87 and F83) can be seen in Table 1. In order to avoid a type II error, additional ANOVAs were made with target words as random sample.

The ANOVA with target word as sample showed that speech rate did in fact have a significant effect on the distance between L and syllable onset (F = 14.095, df 1, p = .013) suggesting that L does not seem to be constantly aligned at or close to the syllable onset. This conclusion is also supported by Figure 6, which, again, displays a large spectrum of alignments both before and after syllable onset. An ANOVA on the distance between H and syllable offset was also calculated which shows a low p-value; however not statistically significant (F = 5.159, df 1, p = .072). Speech rate appears to not affect the possible anchoring of H. Average value and standard deviations for each target word are shown in Table 2.

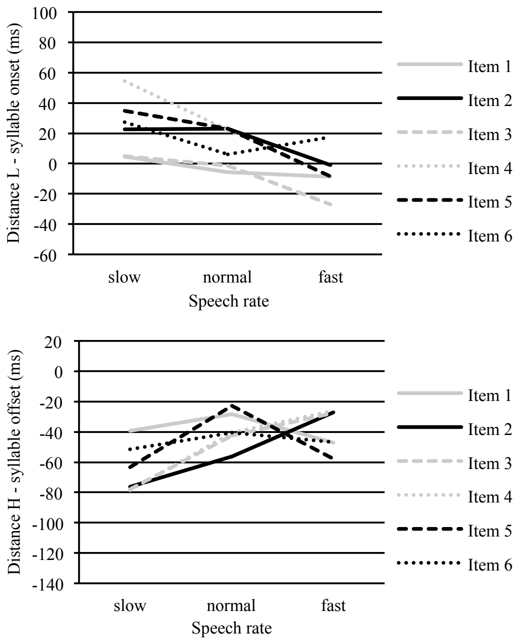

Figure 6: *Distance between tonal target and segment boundary, average value for each target word. The first graph shows distance between L and syllable onset, the second between H and offset. A negative number indicates that timing is before the boundary.*

Table 2: *Distance between tonal target and segment boundary (ms), average value for each target word. Standard deviation in parentheses. A negative number indicates that the target is before the boundary.*

	L - syllable onset			H - syllable offset		
	slow	normal	fast	slow	normal	fast
1	4 (22)	-6 (37)	-9 (37)	-39 (32)	-28 (25)	-47 (37)
2	23 (65)	23 (51)	-1 (18)	-76 (25)	-56 (19)	-27 (37)
3	5 (50)	-2 (59)	-27 (34)	-78 (54)	-42 (19)	-28 (17)
4	54 (56)	22 (78)	-9 (11)	-79 (39)	-40 (21)	-26 (22)
5	35 (50)	23 (59)	-8 (9)	-63 (33)	-23 (17)	-58 (16)
6	27 (64)	6 (60)	18 (59)	-52 (30)	-40 (29)	-47 (18)
All	24 (51)	9 (54)	-6 (36)	-65 (37)	-38 (23)	-40 (27)

4. Discussion

This study seems to not support the anchoring of L in the L+H rise of South Swedish Accent II. The rise is surely an important feature of the word accent [10], but if the start of the rise is not constantly aligned it is possible that L is not a phonological event. The end of the rise, the timing of H, might be an important phonological feature. Independent of syllable duration H was aligned within the syllable, on average 40 ms before syllable offset in the case of normal and fast speech, which supports the Lund Model and the accent typology that incorporates the South Swedish dialect. The study would benefit from relative measures of syllable duration, not the least to further establish the location of H, but also to shed light on the variability of L which affected by speech rate occurs both before and after syllable onset (Figure 5 and 6). Future studies would also benefit from addressing the precision of tonal peak measures (a problem in many comparative studies [1]-[3], [5], [6]), and use alternative methodology, such as Tonal Center of Gravity (TCoG) [12].

The data displayed a great variability both between and within speakers. Ladd et al. [3] also observed a similar degree of variability. By excluding certain speakers that seemed to use a different strategy to define pitch accent, they were able to find support for segmental anchoring. It might be that constant alignment is a strategy only for some speakers or that the same speaker uses different strategies for alignment. The proposition by Niebuhr et al. [7] to include speaker strategy in studies on tonal alignment is thus a valid suggestion.

The auxiliary hypothesis that primary features of phonetic properties will try to be retained by speakers, while other features will be allowed to be modified by time pressure might be the case for the normal and the fast rate. The slow rate, however, seemed to divert from the others, which can be seen in the scatter plot with the slow rate being much more scattered than the normal or the fast rate (Figure 4). The anomalies found on slow speech rate have been reported in other studies as well, where difficulties with the slow speech rate seem to have brought forth additional prosodic features to enable the, perhaps, unnaturally slower speech [3]. Even though the results of the study confirmed that the manipulation of speech rate was successful, a future use of speech rate as an experimental tool needs to be further investigated.

The coincidence was pointed out that the rise of the pre-nuclear accent in intonation languages phonetically roughly matched the lexical Accent II in South Swedish. The results, however, did not confirm a phonological match. The start of a rising pre-nuclear accent in an intonation language needs to be anchored, but this does not seem to be the case for a South Swedish Accent II rise. Since evidence was found against L anchoring with syllable onset, the results do not support the revised Lund Model of a LH gesture. Further studies on the anchoring of the LHL tonal gesture in Swedish Accent II are suggested.

5. Acknowledgements

This paper is based on an unpublished master thesis. While working with an earlier version of the thesis I was supervised by professor Gösta Bruce and associate professor Hugo Quené. I am very grateful for having had their excellent supervision in the initial steps of the study. I also want to extend a big thank you to my supervisor Gilbert Ambrazaitis for his invaluable help, advices and enthusiasm.

6. References

[1] Arvaniti, A., Ladd, D. R. and Mennen, I., "Stability of tonal alignment: the case of Greek prenuclear Accents", Journal of Phonetics, 26:3-25, 1998.

[2] Atterer, M. and Ladd, D. R., "On the Phonetics and Phonology of 'Segmental anchoring' of F0: Evidence from German", Journal of Phonetics, 32:177-197, 2004.

[3] Ladd, D. R., Faulkner, D., Faulkner, H. and Schepman, A., "Constant 'Segmental anchoring' of F0 movements under changes in speech rate", The Journal of the Acoustical Society of America, 106(3):1543-1554, 1999.

[4] Niemann, H., Mücke, D., Nam, H., Goldstein, L. and Grice, M., "Tones as Gestures: the Case of Italian and German", ICPhS XVII Proc., Hong Kong, 1486-1489, 2011.

[5] Caspers, J. and Van Heuven, V. J., "Effects of time pressure on the phonetic realization of the Dutch Accent-lending pitch rise and fall", Phonetica, 50:161-171, 1993.

[6] Xu, Y., "Consistency of tone-syllable alignment across different syllable structures and speaking rates", Phonetica, 55:179-203, 1998.

[7] Niebuhr, O., D'Imperio, M., Fivela, B. G. and Cangemi, F., "Are there 'Shapers' and 'Aligners'? Individual differences in signaling pitch accent category", ICPhS XVII Proc., Hong Kong, 120-123, 2011.

[8] Bruce, G. and Gårding, E., "A prosodic typology of Swedish dialects", in E. Gårding, G. Bruce and R. Bannert [Ed], Nordic prosody, Papers from a symposium, 219-228, Lund, 1978.

[9] Bruce, G., "Components of a prosodic typology of Swedish intonation", in T. Riad and C. Gussenhoven [Ed], Tones and Tunes, Volume 1: Typological Studies in Word and Sentence Prosody, 113-146, Berlin, 2007.

[10] Ambrazaitis, G. and Bruce, G., "Perception of south Swedish word Accents", Working papers, 52:5-8, Lund, 2006.

[11] Boersma, P. and Weenink, D., "Praat: doing phonetics by computer [Computer program]", version 5.2.03. Online: http://www.praat.org/, accessed on 24 November 2010.

[12] Barnes, J., Veilleux, N., Brugos, A. and Shattuck-Hufnagel, S., "Tonal Center of Gravity: A global approach to tonal implementation in a level-based intonational phonology", Laboratory Phonology, 3(2): 337-383, 2012.