Constant tonal alignment in Swedish word accent II

Malin Svensson
Linguistics and Phonetics, Lund University, Lund, Sweden

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.

Introduction

Over a period of 13 years there has been an on-and-off debate within intonational phonology on whether accentual tonal targets (L, H) are constantly aligned with the segmental string. Most studies have focused on pre-nuclear rising accents in intonation languages. A language with lexical accents has not been taken into account in the recent research on constant alignment so far.

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. Studies that second the principle have looked at rising pre-nuclear accents in intonation languages (Arviniti et al., 1998; Atterer & Ladd, 2004; Ladd et al. 1999).

Previous studies in the field 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 (Arvaniti et al., 1998, for Greek; Niemann et al., 2011, for Italian), at syllable onset (Caspers & Van Heuven, 1993, for Dutch; Ladd et al., 1999, for English; Atterer & Ladd, 2004, for German), or after syllable onset (Xu, 1998, for Mandarin).

The precise timing seems to vary, but the studies do show anchoring of L with the beginning of the syllable, while the same consistent result does not exist for the H target. Some studies found H anchoring after syllable offset (Atterer & Ladd, 2004; Xu, 1998; Arvaniti et al., 1998), or somewhere late in the syllable (Ladd et al., 1999). Caspers and Van Heuven (1993) found that the end of the rise, the H-target, varied considerably under time pressure and thus discarded that it did anchor with the segment.

Swedish accents

In the prosodic typology of Swedish intonation provided by the Lund Model (Bruce & Gårding, 1978; Bruce, 2007), the two Swedish word accents are assumed to be represented by a fall in the stressed syllable in a prosodic word, where the two accents differ in the timing of the fall. There is a 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 (2007), 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 (Ambrazaitis & Bruce, 2006).

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 (Caspers & Van Heuven, 1993). Speech rate has been used successfully by a number of researchers in studies concerning tonal alignment (Caspers & Van Heuven, 1993; Ladd et al., 1999; Xu, 1998). Because the Lund Model (revised by Bruce, 2007) 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.

Method

Speakers and recording

The material was initially recorded 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.

Speech materials and data processing

The materials consisted of three test sentences with the same test word. The materials were mixed with 37 further sentences not investigated here. The three test words fit 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 target word *många* [ˈmɔŋːa] 'many' occurred before nuclear accent in all three sentences.

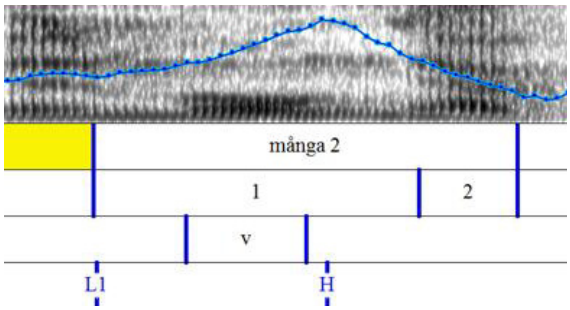

Figure 1. Image of praat window. Target word in sentence 2 Nej, Anna bilar till Polen med många VÄNNER *by speaker M63 in normal speech rate.*

The author performed segmentation and annotation in Praat (Boersma & Weenink, 2010). Since each speaker was recorded twice, the material consisted of 126 items (3 sentences × 2 repetitions × 3 speech rates × 7 speakers). Each target word was segmented into syllables, and in addition, the boundaries of the accented vowel were determined (*Figure 1*). Since the border between the two syllables was difficult to distinguish, offset of accented syllable was calculated to occur in between offset of accented vowel and onset of the following vowel. The tonal curve has been semi-automatically annotated for the tonal targets L and H. Extracted measures were the start and the end of the rise (L and H), syllable onset and syllable offset.

Results

Average syllable duration show a difference in speech rate between the recordings (*Figure 2*). An ANOVA confirmed that speech rate had a significant effect on syllable duration ($F = 66.490$, df 1, $p = .000$), 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 in the segmental string. An ANOVA 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.

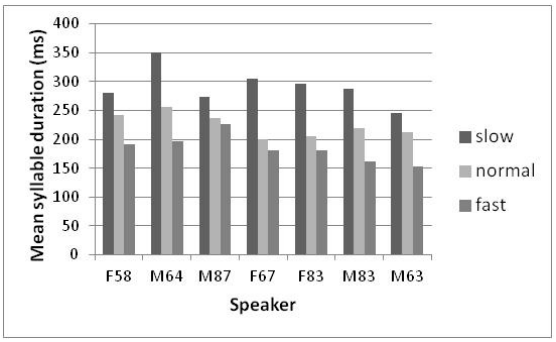

Figure 2. Graph of average syllable duration for each speaker in each speech rate.

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, there does not seem to be a convincingly strong correlation (*Figure 3*).

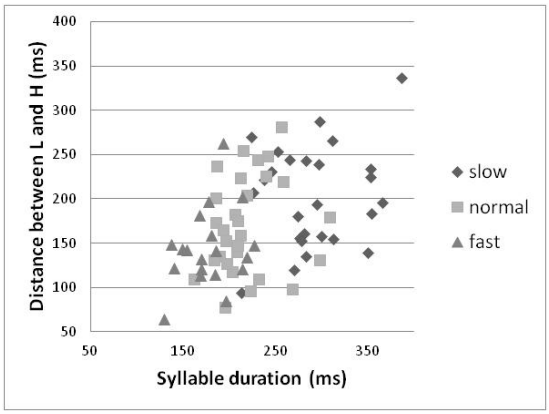

Figure 3. Scatter plot of the relationship between syllable duration and the distance between tonal targets L and H.

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: distance between L and syllable onset, and H and syllable offset. 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, *Table 1* displays quite a large standard deviation as well as the anomaly of only one available item in some conditions for three of the speakers (M64, M87 and F83). To avoid a type II error, additional ANOVAs were made with target words as random sample.

*Table 1. Average distance between tonal target and segment border (ms) for each speaker. Negative number indicate that the target is before the border. Standard deviation in parentheses. * 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 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). This is evidence against segmental anchoring of the L target. 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 and standard deviations are shown in *Table 2*.

Table 2. Average distance between tonal target and segment border (ms) for each target word. Negative number indicate that the target is before the border. Standard deviation in parentheses.

	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)

Discussion

This study does 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 (Ambrazaitis & Bruce, 2006), but if the start of the rise is not anchored 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 the timing of H was found approximately 40 ms before syllable offset, which supports the Lund Model and the accent typology that incorporates the South Swedish dialect.

The data displayed a great variability both between and within speakers. Ladd et al. (1999) 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 of the speakers or that the same speaker might use different strategies for aligning the tones to the segment. The proposition by Niebuhr et al. (2011) to include speaker strategy in investigation 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 seemed, however, to divert from the others and induces other compensatory prosodic features. This can be seen in the scatter plot of the correlation test (*Figure 3*) with the slow rate being much more scattered across the graph than the normal or the fast rate. The anomalies found in this study on slow speech rate have also been reported in other studies, where difficulties with the slow speech rate seem to have brought forth additional prosodic features to enable the, perhaps, unnaturally slower speech (Ladd et al., 1999). 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.

Acknowledgements

While working with an earlier version of this master 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, understanding and enthusiasm.

References

Ambrazaitis, G. & G. Bruce. 2006. Perception of south Swedish word Accents. *Working papers* 52:5–8. Lund University, Lund, Sweden.

Arvaniti, A., D. R. Ladd & I. Mennen. 1998. Stability of tonal alignment: the case of Greek prenuclear Accents. *Journal of Phonetics* 26:3–25

Atterer, M. & D. R. Ladd. 2004. On the Phonetics and Phonology of "Segmental anchoring" of F0: Evidence from German. *Journal of Phonetics* 32:177–197.

Boersma, P. & D. Weenink. 2010. Praat: doing phonetics by computer [Computer program]. Version 5.2.03, retrieved 24 November 2010 from http://www.praat.org/

Bruce, G. 2007. Components of a prosodic typology of Swedish intonation. In Tomas Riad & Carlos Gussenhoven (eds.), *Tones and Tunes, Volume 1: Typological Studies in Word and Sentence Prosody*. Berlin: Mouton de Gruyter, 113–146.

Bruce, G. & E. Gårding. 1978. A prosodic typology of Swedish dialects. In: Eva Gårding, Gösta Bruce & Robert Bannert (eds.) *Nordic prosody. Papers from a symposium.* Lund University, Lund, Sweden, 219–228.

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

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

Niebuhr, O., M. D'Imperio, B. G. Fivela & F. Cangemi. 2011. Are there "Shapers" and "Aligners"? Individual differences in signaling pitch accent category. In: *Proceedings of ICPhS XVII*, Hong Kong, 120–123.

Niemann, H., D. Mücke, H. Nam, L. Goldstein & M. Grice. 2011. Tones as Gestures: the Case of Italian and German. In: *Proceedings of ICPhS XVII*, Hong Kong, 1486–1489.

Xu, Y. 1998. Consistency of tone-syllable alignment across different syllable structures and speaking rates. *Phonetica* 55:179–203.

Proceedings of Fonetik 2013

The XXVIth Annual Phonetics Meeting 12–13 June 2013, Linköping University Linköping, Sweden

Studies in Language and Culture
no. 21

Robert Eklund, editor

Linköping University

Conference website: www.liu.se/ikk/fonetik2013
Proceedings also available at: http://roberteklund.info/conferences/fonetik2013

Cover design and photographs by Robert Eklund
Photo of Claes-Christian Elert taken by Eva Strangert on the occasion of his 80th birthday

Proceedings of Fonetik 2013, the XXVIth Swedish Phonetics Conference
held at Linköping University, 12–13 June 2013
Studies in Language and Culture, no. 21
Editor: Robert Eklund
Department of Culture and Communication
Linköping University
SE-581 83 Linköping, Sweden

ISBN 978-91-7519-582-7
eISBN 978-91-7519-579-7
ISSN 1403-2570

Printed by LiU-Tryck, Linköping, Sweden, 2013