Perception of South Swedish Word Accents

Ambrazaitis, Gilbert; Bruce, Gösta

Published in:
Conference proceedings from Fonetik 2006

2006

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Ambrazaitis, G., & Bruce, G. (2006). Perception of South Swedish Word Accents. In G. Ambrazaitis, & S. Schötz (Eds.), *Conference proceedings from Fonetik 2006: Lund, June 7-9, 2006. Working Papers in Linguistics, Lund University, Centre for Languages and Literature* (Vol. 52, pp. 5-8). (Working Papers in General Linguistics and Phonetics; Vol. 52). Centre for Languages and Literature, Lund University.
http://journals.lub.lu.se/ojs/index.php/LWPL/article/view/2282/1857

Total number of authors:
2

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Working Papers 52 (2006), 5–8

Perception of South Swedish Word Accents

Gilbert Ambrazaitis and Gösta Bruce

Dept. of Linguistics and Phonetics, Centre for Languages and Literature, Lund University
{Gilbert.Ambrazaitis|Gosta.Bruce}@ling.lu.se

Abstract

A perceptual experiment concerning South Swedish word accents (accent I, accent II) is described. By means of editing and resynthesis techniques the F0 pattern of a test word in a phrase context has been systematically manipulated: initial rise (glide vs. jump) and final concatenation (6 timing degrees of the accentual fall). The results indicate that both a gliding rise and a late fall seem necessary for the perception of accent II, while there appear to be no such specific, necessary cues for the perception of accent I.

1 Introduction

In the original Swedish intonation model (Bruce & Gårding, 1978) the two tonal word accents (accent I and accent II) are assigned bitonal representations in terms of High plus Low (HL), representing the accentual F0 fall. These Highs and Lows are timed differently, however, in relation to the stressed syllable depending on dialect type. For all dialect types, the HL of accent I precedes the HL of accent II. In South Swedish, the HL of accent I is aligned with the stressed syllable, while the HL of accent II is instead aligned with the post-stress syllable.

A problem with the latter representation is that the stressed syllable in accent II words has no direct tonal representation. Thus this modelling does not reflect what should be the most perceptually salient part of the pitch pattern of accent II. Figure 1 shows prototypical F0 contours of the two word accents (minimal pair) in a prominent position of an utterance as produced by a male speaker of South Swedish (the second author).

This particular problem of intonational modelling has been the starting-point of a phonetic experiment aimed at examining what is perceptually relevant in the F0 contours of accent I and accent II in the South Swedish dialect type. More specifically, our plan has been to run a perceptual experiment, where the intention was to find out what are the necessary and sufficient cues for the identification of both word accents.

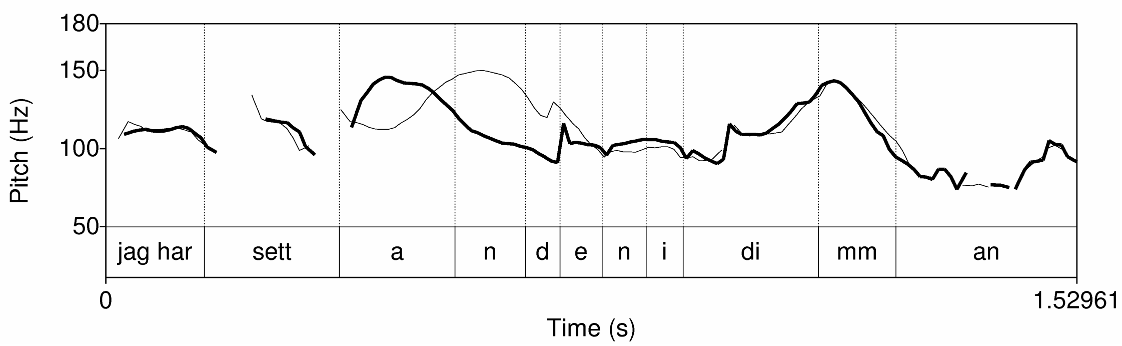

Figure 1. Prototypical F0 contours of the two word accents in a prominent position of an utterance as produced by a male speaker of South Swedish: *Jag har sett anden i dimman. ('I have seen the duck/spirit in the fog.')* Thick line: acc. I *('duck')*; thin line: acc. II *('spirit')*.

2 Method

We asked subjects to judge whether they perceive the test word *anden* as either meaning *'the duck'* (accent I) or *'the spirit'* (accent II), in naturally produced and synthetically manipulated test utterances. We chose to put the test word in a non-final accented position of an utterance containing two accented words (test word and context word; see Table 1), for several reasons. First, we wanted to have the possibility of removing the accentual F0 fall of the test word while maintaining an utterance-final falling pattern. Second, we chose two different context words – one with accent I (*drömmen, 'the dream'*), one with accent II (*dimman, 'the fog'*) – in order to provide a "dialectal anchor" for the listeners. Third, by having the test word non-finally, we avoided phrase-final creaky voice on the test word, thus facilitating the editing of F0. Regarding semantic factors, we tried to choose context words which would be as "neutral" as possible, i.e. which would not bias the ratings of the test word. The test material was recorded by a male speaker of South Swedish (the second author) in the anechoic chamber at the Centre for Languages and Literature, Lund University.

Table 1. The structure of the test material, or the four recorded test utterances respectively. (*'I have seen the duck/spirit in the dream/fog.'*)

	Test word	Context word	Used for
Jag har sett	*anden (**accI**)*	*i drömmen (**accI**) / i dimman (**accII**).*	A: control stimuli
	*anden (**accII**)*	*i drömmen (**accI**) / i dimman (**accII**).*	B: primary stimuli

2.1 Stimuli

We created 12 F0 contours and implemented them in two recorded utterances (B in Table 1), by means of F0 editing and resynthesis using the PSOLA manipulation option in Praat (<http://www.praat.org>). Figure 2 displays the 12 contours for one of the utterances (*dimman*) as an example. The starting point was a stylization of the originally recorded F0 contours, i.e. with accent II on the test word (glide/dip4 in Figure 2). Based on this stylized accent II contour, three contours with a successively later F0 fall were created (dip3, dip2, dip1), each one aligned at successive segmental boundaries: in dip3, the fall starts at the vowel onset of the post-stress syllable (schwa), in dip2 at the following /n/ onset, and in dip1 at the onset of /i/. Thus, a continuum of concatenations between the two accented words was created. Two further steps were added to this continuum: one by completely removing the fall, yielding a contour that exhibits a high plateau between the two accented words (dip0), and one by shifting back the whole rise-fall pattern of the original accent II, yielding a typical accent I pattern (dip5). For each dip position, we also created a contour that lacks the initial

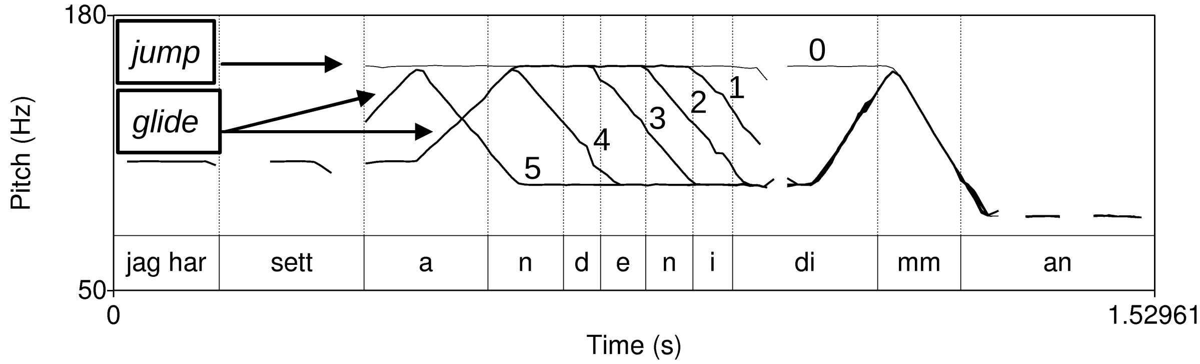

Figure 2. Stimulus structure, exemplified in the *dimman* context: 6 dip levels (0...5) x 2 rise types (jump, glide). These 12 F0 contours were implemented in both recordings (*dimman* and *drömmen*), yielding 24 stimuli.

gliding rise on /a(n)/, by simply transforming it into a “jump” from low F0 in *sett* to high F0 right at the onset of *anden*. It should be pointed out that the difference between glide and jump is marginal for dip5 (i.e. accent I), and was implemented for the sake of symmetry only.

Additionally, we generated 4 control stimuli which were based on the A-recordings (cf. Table 1). These are, however, not further considered in this paper.

2.2 Procedure

All 24+4=28 stimuli were rated 4 times. The whole list of 112 stimuli was randomized and presented to the listeners in 8 blocks of 14 stimuli each, via headphones. The listeners heard each stimulus only once and had to rate it as either referring to a duck (*and*) or a spirit (*ande*), within 3 seconds, by marking it on a paper sheet. The whole test was included in a wav-file and took 11:31 minutes. Instructions were given orally and in written form. A training test with two blocks of 4 stimuli each was run before the actual experiment. 20 South Swedish native speakers, 5 male, 15 female, aged 19-32, with no reported hearing impairments, volunteered as subjects.

2.3 Data analysis

Based on the four repetitions of each stimulus, an accent II score in % (henceforth %accII) was calculated per stimulus and subject. These %accII scores were used as raw data in the analyses. Means and standard deviations, pooled over all 20 listeners, were calculated for every stimulus. A three-way repeated-measures ANOVA was run for the 24 primary stimuli to test for effects of the following factors: FINAL WORD (2 levels: drömmen, dimman), RISE TYPE (2 levels: jump, glide), and CONCATENATION (6 levels: dip0…dip5).

3 Results

The mean %accII ratings are displayed in Figure 3. The stimuli that were intended to represent clear cases of accent I (dip5), and accent II (glide/dip4) were convincingly rated as expected. The graphs for the two different contexts look very similar, and accordingly, FINAL WORD had no significant effect ($p>.8$). Also, as would be expected from Figure 3, both RISE TYPE and CONCATENATION have a significant effect ($p<.001$ each). However, the difference in rise type is not reflected in a constant %accII difference, which is especially salient in dip5. Accordingly, we also found a significant interaction between RISE TYPE and CONCATENATION ($p<.001$).

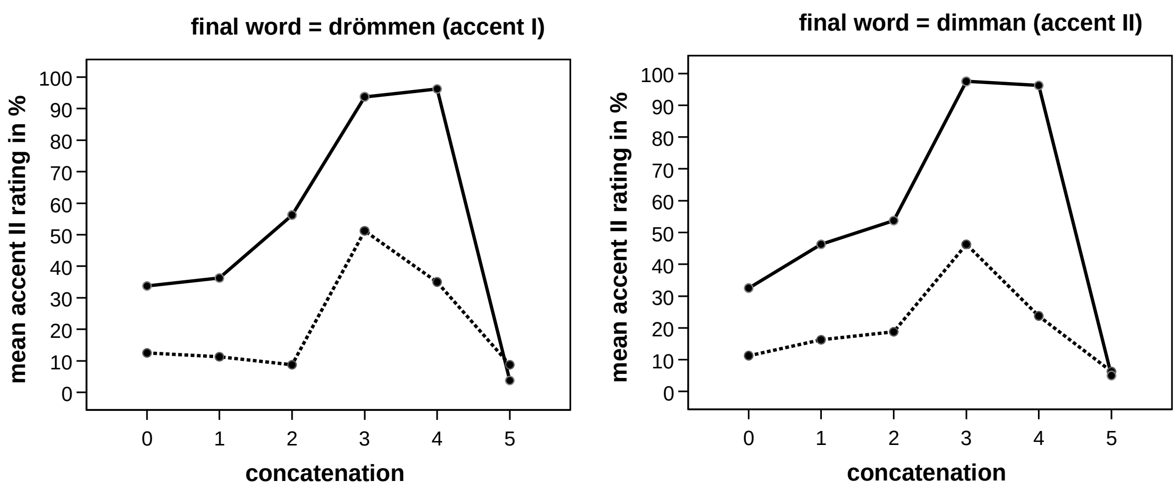

Figure 3. Mean accII-ratings in % for 2 final word conditions, 6 dip levels (concatenation), and 2 rise types: glide (straight line) and jump (dotted line).

4 Discussion

Referring back to the issue about necessary and sufficient cues for word accent identification (cf. Introduction), we will comment on a number of points in the light of our experiment.

Is the gliding rise through the stressed syllable necessary for the perception of accent II? – Replacing this glide by an F0 jump up to the stressed syllable results in a sizeable decrease in the votes for accent II (cf. glide/dip4 vs. jump/dip4). This suggests that the gliding rise is necessary for the unambiguous perception of accent II.

Is a late-timed fall necessary for the perception of accent II? – Replacing the F0 fall through the post-stress syllable by a high plateau in the target word yields a tendency towards accent I (cf. glide/dip4 vs. glide/dip0). This suggests that the fall is necessary. However, the fall must not be substancially earlier than in the original accent II word, since this would correspond to accent I.

Thus, both the gliding rise and the late fall seem necessary for the unambiguous perception of accent II. When one of them is removed, the ratings tend towards accent I. When both these cues are absent, the tendency becomes rather strong (cf. jump/dip0).

What is necessary, and what is sufficient for the perception of accent I? – Accent I is most convincingly represented by stimuli with an early fall (dip5). However, the discussion above has already shown that this early fall cannot be a necessary cue, since a number of stimuli lacking this early fall have received high accent I ratings (cf. also jump/dip1-2). Furthermore, a high-starting stressed syllable (jump) favors accent I ratings, but cannot be regarded as sufficient, since the absence of an accent II-like fall appears necessary (cf. dip3).

Thus, our results are most easily explainable along the hypothesis that there are no specific necessary cues for accent I at all, but that simply the absence of accent II cues is sufficient for the perception of accent I. It is still remarkable, though, that the absence of only one accent II cue alone (e.g. the late fall) results in more votes for accent I than for accent II.

Why does a glide followed by a plateau trigger a considerable number of votes for accent I? – We do not have a definite answer to this question. One possibility is that the conditions of phrase intonation play a role. In the early part of a phrase, the expectation is a rising pattern. Thus a phrase-initial accent I may be realized as a rising glide even in South Swedish, as long as there is no immediately following F0 fall. Another possibility is that the glide-plateau gesture represents a typical accent I pattern of another dialect type (Svea or Stockholm Swedish), even if the context word at the end has a stable South Swedish pattern.

What does our experiment tell us about the markedness issue? – From the perspective of perceptual cues, accent II in South Swedish appears to be more “special” than accent I. This will lend some support to the traditional view of accent II being the marked member of the opposition (cf. Elert, 1964; Engstrand,1995; Riad 1998).

Acknowledgements

Joost van de Weijer assisted us with advice concerning methodology and statistics.

References

Bruce, G. & E. Gårding, 1978. A prosodic typology for Swedish dialects. In E. Gårding et al. (eds.), *Nordic Prosody*. Lund: Lund University, Department of Linguistics, 219-228.

Elert, C-C., 1964. *Phonologic Studies of Quantity in Swedish*. Stockholm: Almqvist & Wiksell.

Engstrand, O., 1995. Phonetic interpretation of the word accent contrast in Swedish. *Phonetica 52*, 171-179.

Riad, T., 1998. Towards a Scandinavian accent typology. In W. Kehrein & R. Wiese (eds.), *Phonology and Morphology of the Germanic Languages*. Tübingen: Niemeyer, 77-109.