

What affects recognition most – wrong word stress or wrong word accent?

Åsa Abelin¹ and Bosse Thorén²

¹Department of Philosophy, Linguistics and Theory of Science, University of Gothenburg

²School of Humanities and Media Studies, Dalarna University

Abstract

In an attempt to find out which of the two Swedish prosodic contrasts of 1) word stress pattern and 2) tonal word accent category has the greatest communicative weight, a lexical decision experiment was conducted: in one part word stress pattern was changed from trochaic to iambic, and in the other part trochaic accent II words were changed to accent I.

Native Swedish listeners were asked to decide whether the distorted words were real words or 'non-words'. A clear tendency is that listeners preferred to give more 'non-word' responses when the stress pattern was shifted, compared to when word accent category was shifted. This could have implications for priority of phonological features when teaching Swedish as a second language.

Introduction

This study started with a discussion at the Phonetics meeting 2014. The topic concerned Swedish spoken with a foreign accent, and whether wrong word stress or wrong word accent was most detrimental for recognition and understanding of words. When teachers make curricula for second language speakers, they are helped by knowing which phonetic or phonological features are more or less crucial for the understanding of speech. There are some structural and anecdotal evidence that word stress should play a more important role in the perception and understanding of Swedish than the tonal word accent. The aim of the study is to find out which of two distortions causes the most difficulty in identifying some disyllabic words: 1) changing the word stress category from trochaic to iambic or 2) changing the tonal word accent category from accent II to accent I.

Background

Swedish word stress is about prominence contrasts between syllables, mainly signalled by syllable duration (Fant & Kruckenberg 1994), although F0 gestures, voice source parameters and differences in vowel quality combine to signal syllable prominence (ibid.) The tonal word accent, however, is mainly signalled by changes in the F0 curve and the timing of those changes within the word. According to Bruce (1977, 2012) and Elert (1970), word stress in Swedish is

variable, and words can have different meanings depending on where the main stress is placed, as found in 'banan' 'the path/course' and *ba'nan* 'banana'. A great number of disyllabic trochaic-iambic minimal pairs can be created. A smaller number of trisyllabic minimal pairs, such as 'Israel' 'the state of Israel' and *isra'el* 'Israeli citizen', are also possible.

According to standard accounts Swedish has two word accent categories, accent I (acute), e.g. *tómtén* 'the plot', and accent II (grave), *tomten* 'Santa Claus', cf. Elert (1970), even though only the grave accent can be considered a real word accent. It is the only one of these two that predicts that the main stressed syllable and the following syllable belong to the same word (in a two-syllable word) i.e. having a cohesive function, and it is limited to the word, simple or compound. The word accent is connected with a primary stressed syllable. In isolation the words usually carry sentence accent and accent II then tends to involve two F0 peaks.

Method

Material and design

The material consisted of 10 trochaic (accent I) words, e.g. *bílen* 'the car', 10 originally trochaic words pronounced with iambic stress, e.g. *vägén* 'the road', 10 iambic words, e.g. *kalás* 'the party', 10 accent II words, e.g. *gatan* 'the street', 10 originally Accent II words pronounced with trochaic stress accent I, e.g. *sagan* 'the fairy

tale', and finally 26 disyllabic non-words, with varying stress or tonal accents. All the words were nouns in the definite form (with one exception) apart from the iambic words. The words were recorded by a male phonetician with a neutral dialect. Recording and editing was made with the software Praat (Boersma & Weenink, 2013).

There was some deliberation about how to treat vowel quality in the stressed and unstressed syllables, since these vary according to degree of stress. We decided to choose vowels which do not vary so much in unstressed vs. stressed position, e.g. /e/ rather than /a/, and keep the quality of the original word, e.g. not changing [e] to [ɛ] in unstressed position. Each word was presented until it self-terminated, in all cases below 1000 ms. Simultaneously the subjects had 1000 ms to react to each stimulus. The time allotted for reaction to the stimuli thus started when the word started. Between each word there was a 1000 ms pause.

For building and running the experiment the software PsyScope was used (Cohen, MacWhinney, Flatt & Provost, 1993).

Procedure

A lexical decision test was performed where 18 female L1 speakers of Swedish, approximately

word was a non-word. The subjects were instructed to decide as quickly as possible, whether the word they heard was a real word or not. Reactions that were not registered within the 1000 ms period were categorized as loss.

Results

Accuracy

Figure 1 shows the main results of the experiment. It turned out that the task was quite difficult, and that the loss in the experiment was large.

Loss signifies that the reaction times were too long, over 1000 ms. The difficulty of responding quickly could be due to the fact that the word stimuli were quite long, between 660 and 979 ms. However, the main result does not concern the reaction times, but the difference in assessment of word status of the words with wrongly pronounced tonal word accent, compared with the words with wrongly pronounced stress placement.

Figure 1 shows that the mispronounced words that were generally judged as real words were the accent II words pronounced with accent I (23% 'yes' responses and 3% 'no' responses), while the words that were generally judged as non-words were the trochaic accent I words

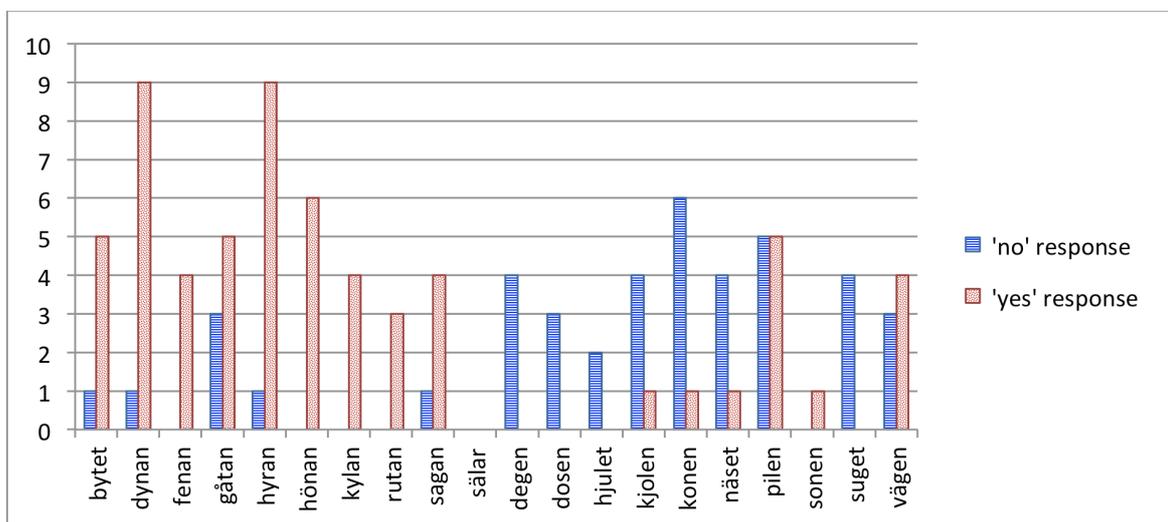


Figure 1. Number of persons deciding on mispronounced words being real words (dotted bars) or non-words (striped bars). The first 10 words to the left are accent II words pronounced with accent I, and the next 10 words, to the right, are trochaic accent I words pronounced with iambic accent.

20–25 years of age, heard the above described 76 words, one by one in random order. The subjects were instructed to press one key on a keyboard if the word was a real word and another key if the

pronounced as iambic accent I (7% 'yes' responses and 19% 'no' responses). This suggests that identification and comprehensibility of speech is more affected by wrong stress

placement in comparison with wrong word accent. Furthermore, there was a larger loss for the words with wrong stress placement than for the words with wrong tonal word accent, however not significant.

Figure 2 compares the wrongly pronounced words with the correctly pronounced words. The figure shows that the correctly pronounced words are the most robust; they exhibit a smaller loss and they are more often assessed as real words. The words which were most frequently misjudged were the words with wrong stress placement.

There is interaction between loss, 'no' responses and 'yes' response: Where there are more 'no' responses the loss is greater. This could be due to the simple fact that 'no' responses generally have longer reaction times than 'yes' responses; thus, it could be that in some cases when a 'no' response is intended the response time exceeds 1000 ms. But the result could also be due to an impossibility to interpret the wrongly pronounced word.

which are generally longer. To compare reaction times for the 'yes' responses is not relevant since there were so few 'yes' responses for the words with wrong stress placement.

Durations of sound stimuli

The durations of the sound stimuli were measured, and we found that the wrongly pronounced trochaic accent I words, pronounced as iambic, were slightly longer. However this did not correlate with reaction times.

In general, reaction times were longer than the word durations, but not if deducting 200 ms for motor action. There is a tendency that when the durations are shorter, loss is smaller and the 'yes' responses are more numerous.

Discussion and conclusion

The results can be discussed in relation to "left-to-right" models of speech perception and to where the actual recognition point is situated (cf.

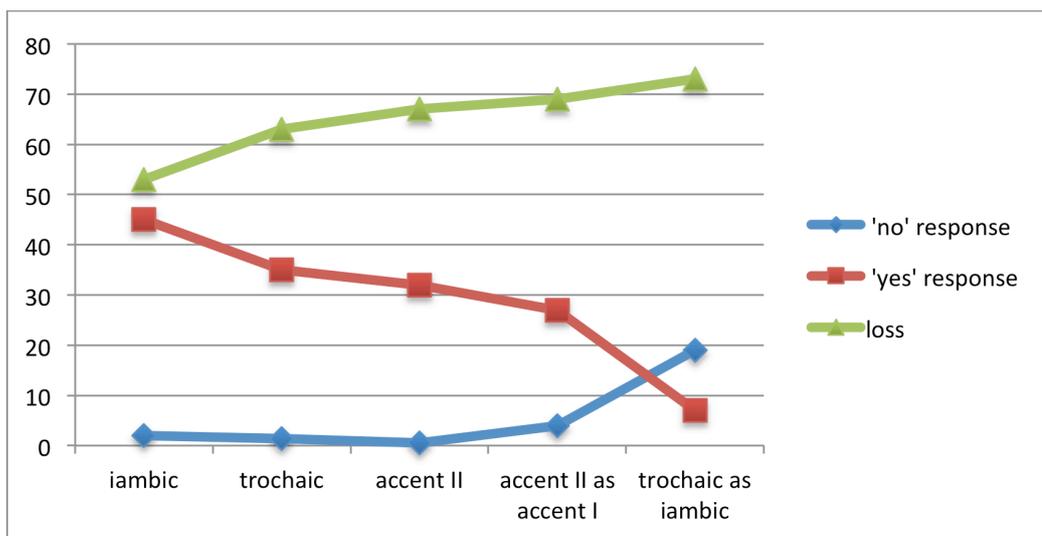


Figure 2. Mean percent for decisions on mispronounced words being real words (squares) or non-words (diamonds), in comparison with correctly pronounced words; iambic, trochaic and accent 2. Triangles stand for loss.

Reaction times

There was not a large difference in mean reaction time between the two wrongly pronounced groups. Mean value for mispronounced accent II was 877 ms and mean value of mispronounced trochaic accent I was 915 m sec. This is not a significant difference, but it is nevertheless like comparing apples with pears, i.e. actually comparing 'yes' responses with 'no' responses,

Marslen-Wilson, 1987). Is prosody a factor here alongside segmental information? One question is whether an early absence of stress placement would be more detrimental for recognition than a late absence, i.e. would a stress-placement-changed trochaic word (which ought to have stress on the first syllable) be more difficult to process than a stress-placement-changed iambic word (which ought to have stress on the second

syllable)? And similarly, would accent I words pronounced with accent II be more difficult to recognize than accent II words pronounced with accent I? Such experiments are presently carried out and analyzed. Preliminary results show the same general tendency as in the present experiment; misplaced stress is more detrimental to recognition and yields more ‘no’ responses than words pronounced with the wrong tonal word accent. There is also a tendency for longer reaction times for misplaced stress compared with mispronounced word accent.

The words of the present experiment were not checked for frequency or number of phonological neighbours. It could be the case that some of the iambic words (which often are loan words) have a lower frequency. On the other hand, the correctly pronounced iambic words were the words that had the least loss, the highest number of ‘yes’ responses and the lowest number of ‘no’ responses, which might indicate an effect of few phonological neighbours, as concerns “stress related neighbours”. The reason that words were not balanced for frequency was that it was difficult to find suitable words. However, frequency is not a main issue since the results mainly concern correct interpretation or misinterpretation, not reaction time.

Another reflection is the following: What does it entail that the iambic (correct) words are not in the definite form? Morphology, such as different inflectional forms, can affect processing. Söderström (2012) studied perception of accent I and accent II in a mismatch condition where accent I words were followed by accent II inducing suffixes, and accent II words were followed by accent I inducing suffixes. He found that there is a stronger relation between suffixes and accent II compared with accent I, which could imply that accent II could indeed be very important to perception, identification and comprehension in certain contexts.

In relation to the studies of Söderström (2012), Söderström, Roll & Horne (2012) the question arises whether accent II might be more important to comprehension where there are

other errors, e.g. in the speech of learners of Swedish as a second language, which might use the wrong suffixes on nouns or verbs. Adding further learner errors such as word order mistakes or wrong lexical choices the picture becomes complicated.

We are well aware of that our experiment does not show high ecological validity since it tested deliberately mispronounced words which were judged out of context. Follow-up studies will hopefully be made in more natural scenarios.

However, the present results suggest that learners of Swedish as a second language benefit more from proficiency in stress placement than in choice of word accent category or precise realization of word accent category.

This is also indicated by the fact that word accent categories are realized differently in different geographical regions, and that some varieties do not utilize the contrast at all.

References

- Boersma, P. & Weenink, D. (2013). Praat: Doing phonetics by computer (<http://www.praat.org>).
- Bruce, G. (1977). *Swedish word accents in sentence perspective* (Vol. 12). Lund University.
- Bruce, G. (2012). *Allmän och svensk prosodi*, Studentlitteratur, Lund.
- Cohen J.D., MacWhinney B., Flatt M., & Provost J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, 25(2), 257-271.
- Elert, C.-C. (1970). *Ljud och ord i svenskan*, Stockholm: Almqvist & Wiksell.
- Fant, G. & Kruckenberg, A. (1994). Notes on stress and word accent in Swedish *STLQPSR*. 2-3.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25: 71–102.
- Söderström, P. (2012). *Processing Swedish word accents - evidence from response and reaction times*, MA thesis in General linguistics, Lund University.
- Söderström, P., Roll, M., & Horne, M., (2012). Processing morphologically conditioned word accents. *Mental Lexicon* 7, 77–89.
- Thorén, B. (2008). *The priority of temporal aspects in L2-Swedish prosody: Studies in perception and production*. PhD thesis, Stockholm University.