The relative perceptual weight of two Swedish prosodic contrasts

Åsa Abelin[1], Bosse Thorén[2]
asa.abelin@ling.gu.se, bth@du.se

[1]Department of Philosophy, University of Gothenburg
[2]School of Humanities and Media Studies, Dalarna University

Abstract. In addition to 9 vowel and 18 consonant phonemes, Swedish has three prosodic phonemic contrasts: word stress, quantity and tonal word accent. There are also examples of distinctive phrase or sentence stress, where a verb can be followed by either an unstressed preposition or a stressed particle. This study focuses on word level and more specifically on word stress and tonal word accent in disyllabic words. When making curriculums for second language learners, teachers are helped by knowing which phonetic or phonological features are more or less crucial for the intelligibility of speech and there are some structural and anecdotal evidence that word stress should play a more important role for intelligibility of Swedish, than the tonal word accent. The Swedish word stress is about prominence contrasts between syllables, mainly signaled by syllable duration, while the tonal word accent is signaled mainly by pitch contour. The word stress contrast, as in *armen* [ˊarːmən] 'the arm' - *armén* [arˊmeːn] 'the army', the first word trochaic and the second iambic, is present in all regional varieties of Swedish, and realized with roughly the same acoustic cues, while the tonal word accent, as in *anden* [ˊanːdən] 'the duck' - *anden* [ˋanːdən] 'the spirit' is absent in some dialects (as well as in singing), and also signaled with a variety of tonal patterns depending on region. The present study aims at comparing the respective perceptual weight of the two mentioned contrasts. Two lexical decision tests were carried out where in total 34 native Swedish listeners should decide whether a stimulus was a real word or a non-word. Real words of all mentioned categories were mixed with nonsense words and words that were mispronounced with opposite stress pattern or opposite tonal word accent category. The results show that distorted word stress caused more non-word judgments and more loss, than distorted word accent. Our conclusion is that intelligibility of Swedish is more sensitive to distorted word stress pattern than to distorted tonal word accent pattern. This is in compliance with the structural arguments presented above, and also with our own intuition.

Keywords: second language pronunciation, intelligibility, word stress, tonal word accent

Introduction

In the field of second language teaching, there are four main skills that normally are considered; listening comprehension, reading comprehension, oral proficiency and writing proficiency. Oral proficiency can be further divided into pragmatics, like turn-taking, fluency and pronunciation. Pronunciation can be divided into segmental – including phonotactics – and prosodic features. Finally, prosodic features can be divided into dynamic, temporal and tonal variables. This study looks particularly at the perceptual weight of temporal vs tonal prosodic features in Swedish. The result could provide some guidelines as to what phonological features could be given higher or lower priority when Swedish is taught as a second language. This paper reports an expanded version of our experiment presented at Fonetik 2015 (Abelin & Thorén, 2015)

According to Munro and Derwing (1995) a foreign accent per se decreases intelligibility to some degree, but increased perceived degree of foreign accent does not seem to reduce intelligibility. We believe however, that specific details in a foreign accent may be more crucial to intelligibility than the perceived degree of global foreign accent. For English, some 'Lingua Franca Core' features were suggested by Jenkins (2002), and for Swedish Bannert (1980) suggested that some phonological features were more crucial to intelligibility than others. Thorén (2008) discussed differentiated priority among Swedish prosodic contrasts and their respective acoustic correlates.

Standard Swedish has three prosodic phonological contrasts: stress placement, quantity and a tonal word accent. There is some structural and anecdotal evidence that word stress should play a more

important role in the perception and understanding of Swedish, than tonal word accent. Henceforth we will discuss only the two latter contrasts. Although both contrasts are phonemic, some dialects like standard Finland-Swedish lack the tonal word accent contrast but are still easily understood by speakers of other regional varieties. Also, in singing the tonal word accent is totally neutralized. 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 and vice versa, or 2) changing the tonal word accent category from accent II to accent I and vice versa.

Swedish word stress is about prominence contrasts between syllables, mainly signaled by syllable duration (Fant & Kruckenberg, 1994), although F0 gestures, voice source parameters and differences in vowel quality combine to signal syllable prominence (ibid.). Tonal word accent, however, is mainly signaled 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* [ˋbɑːnan] 'the path/course' and *banan* [baˊnɑːn] 'banana'. A great number of disyllabic trochaic-iambic minimal pairs can be created. A smaller number of trisyllabic minimal pairs, such as *Israel* [ˋiːsrael] 'the state of Israel' and *israel* [ɪsraˊeːl] 'Israeli citizen', are also possible.

According to standard accounts Swedish has two word accent categories: accent I (acute), as in *tomten* [ˊtɔmːtən] 'the plot', and accent II (grave), as in *tomten* [ˋtɔmːtən] 'Santa Claus' (see 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 disyllable 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. Pronounced in isolation, words usually carry sentence accent and accent II then tends to involve two F0 peaks.

The purpose was thus to investigate the relative perceptual weights of the two prosodic contrasts, and the weight of the categories of each contrast. The purpose of the first experiment was to test the recognition of words with trochaic stress mispronounced with iambic stress, and words with accent II mispronounced with accent I. The purpose of the second experiment was to test the recognition of words with iambic stress mispronounced with trochaic stress, and words with accent I mispronounced with accent II.

Method

Material and design

The material for the first experiment consisted of 10 trochaic (accent I) words, e.g., *bilen* [ˊbiːlen] 'the car', 10 originally trochaic words pronounced with iambic stress, e.g., *vägen* *[vɛˈgɛn] 'the road', 10 iambic words, e.g., *kalas* [kaˊlɑːs] 'the party', 10 accent II words, e.g., *gatan* [ˋgɑːtan] 'the street', 10 originally Accent II words pronounced with trochaic stress and accent I, e.g., *sagan* *[ˊsɑːgan] 'the fairy tale', and finally 26 disyllabic non-words with varying stress or tonal accents. Furthermore, the material for the second experiment consisted of 10 trochaic (accent I) words, e.g., *köket* [ˊɕøːkət] 'the kitchen', 10 originally iambic words pronounced with trochaic stress, e.g., *kanel*, *[ˊkaneːl] 'cinnamon', 10 iambic words, e.g., *kalas* [kaˊlɑːs] 'the party', 10 accent II words, e.g., *gatan* [ˋgɑːtan] 'the street', 10 originally accent I words pronounced with accent II, e.g., *djuret* *[ˋjʉːrət] 'the animal'. The same 26 disyllabic non-words as in the first experiment were used.

All trochaic words (with one exception) were nouns in the definite form. The words were recorded by a male phonetician with a neutral dialect. Recordings were made with a Røde NT3 condenser microphone to a laptop in a silent studio in the University of Umeå, Sweden, and editing was made with the Praat software (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] or [ə] in unstressed position. Each word was presented until it self-

terminated, in all cases just 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 PsyScope software was used (Cohen, MacWhinney, Flatt, & Provost, 1993).

Procedure

Two lexical decision tests were performed. In the first experiment there were 18 female L1 speakers of Swedish, approximately 20–25 years of age, who were presented with the above described 76 words of experiment 1, one by one in random order. In the second experiment, there were 16 female L1 speakers of Swedish, approximately 20–25 years of age, who were presented with the above described 76 words of experiment 2, 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 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. The subjects had no reported hearing impairment.

Results

Accuracy

Figure 1 shows the main results of experiments 1 and 2. It turned out that the task was quite difficult, and that the loss in the experiment was large.

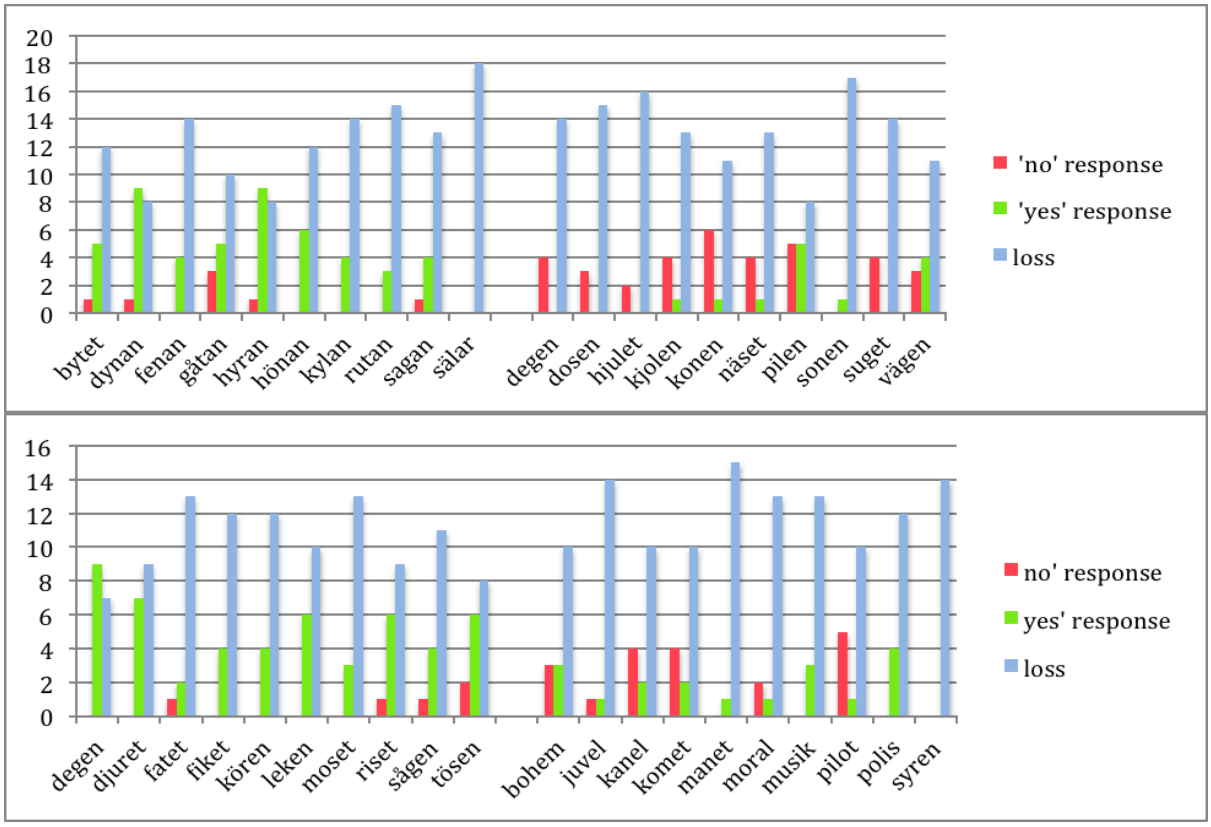

Figure 1. Main results of experiment 1 (above) and experiment 2 (below). The ten bars to the left show the effect of wrong tonal word accent, while the ten bars to the right show the effect of wrong stress placement.

It is evident from Figure 1 that wrong stress placement produced more rejections than wrong tonal word accent in both experiments.

Wrong tonal accent produced more acceptance than wrong stress placement in both experiments. An unpaired t-test showed a significant difference between the two groups ($p < .0001$). The difference in number of 'yes' responses between accent I mispronounced as accent II and accent II mispronounced as accent I is not significant. Neither is the difference between trochaic as iambic and iambic as trochaic significant.

Figure 2 shows a comparison between the wrongly pronounced words with the correctly pronounced words. The figure shows that the correctly pronounced words are, as expected, the most robust; they exhibit a smaller loss and they are more often assessed as real words. The words, which were most frequently judged as non-words were the words with wrong stress placement. The difference in number of 'yes' responses between correctly pronounced accent I words and accent I words pronounced with accent II was significant in an unpaired t-test ($p = .0233$). The difference in number of 'yes' responses between correctly pronounced accent II words and accent II words pronounced with accent I was not significant. When comparing the numbers for loss, accent II pronounced as accent I showed a larger loss than the reverse condition.

The difference in number of 'yes' responses between correctly pronounced trochaic words and trochaic words pronounced with iambic stress was significant ($p<.0001$). Likewise, the difference in number of 'yes' responses between correctly pronounced iambic words and iambic words pronounced with trochaic stress was significant ($p<.0001$).

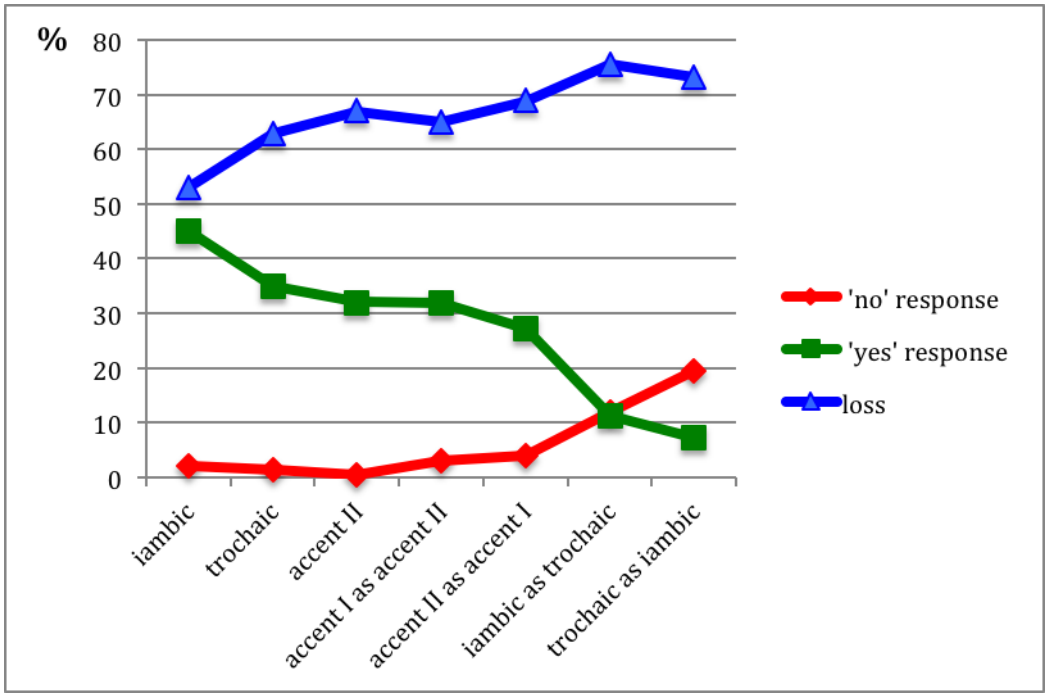

Figure 2. Results for Experiments 1 and 2, including the correctly pronounced words, shown to the left in the diagram.

There is interaction between loss, 'no' responses and 'yes' responses. There is a negative correlation between number of 'yes' responses and loss ($r^2 = .8473$). Furthermore, where there are more 'no' responses the loss is greater.

Reaction times

There was no large reaction time difference in mean total between the wrongly pronounced groups. However, to compare reaction times for the 'yes' responses is not possible 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 activation. There is a tendency that when durations are shorter, loss is smaller and the 'yes' responses are more numerous.

Further experiments

We are also performing an experiment with longer reaction times in order to see what happens to the variable loss. Preliminary results show that loss is diminished when longer reaction times are allowed. Reaction times for 'yes' responses are now possible to measure, and the length of reaction times partly reflects the same order as shown in diagram 2: trochaic accent I words have the shortest reaction times, thereafter iambic words and then trochaic accent II words. Furthermore, accent II mispronounced as accent I have longer reaction times than the correctly pronounced words, and the longest reaction times are the responses to iambic words pronounced as trochaic. Also here we can preliminarily conclude that wrong stress placement is more detrimental to identification of words, since wrong stress placement produces longer reactions times.

We plan to undertake further experiments for testing mispronunciations of the quantity contrast, such as the one in the minimal pair *vila* [ˋviːla] 'to rest' and *villa* [ˋvɪlːa] 'villa'. Together with the present results, the quantity data could help provide a more complete ranking among the Swedish prosodic contrasts with respect to their importance for communication and education.

Discussion

The results suggest a greater perceptual weight for stress pattern when compared with tonal word accent. Furthermore, the results can be discussed in relation to "left-to-right" models of speech perception and to where the actual recognition point is (cf. Marslen-Wilson, 1987). 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)? There is some evidence for this, although the difference was not significant: wrongly pronounced trochaic words were more difficult to identify than wrongly pronounced iambic words. Thus, an early absence of stress placement is more difficult to process.

The words of the present experiments were not checked for frequency or number of phonological neighbors. It could be the case that some of the iambic words (which often are loan words) have a lower frequency. On the other hand, correctly pronounced iambic words were 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 neighbors, as concerns "stress related neighbors". The reason that words were not balanced for frequency was that it was difficult to find suitable words. We made a check for possible correlations between rankings of frequencies and rankings of reaction times, and found no correlation between lower frequencies and longer reaction times. 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) and Söderström, Roll, and 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. When adding further learner errors such as word order mistakes or wrong lexical choices, the picture becomes complicated.

There was an interaction between loss and 'yes' responses, where there was a negative correlation between number of 'yes' responses and loss. Furthermore, where there were more 'no' responses, the loss was 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. This is further explored in an experiment with the possibility for longer reaction times.

We are well aware that our experiment does not show high ecological validity since it tested deliberately mispronounced words that 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.

Conclusion

We conclude that Swedish L1 listeners perceive and identify words with incorrect stress placement and incorrect tonal word accent with greater difficulty than words pronounced with correct stress and correct word accent. Thus, correctly pronounced words were easier to identify (they produced smaller loss, more 'yes' responses and less 'no' responses) than the wrongly pronounced words. There was a difference in the order of ease for identifying the correctly pronounced words: the easiest were the iambic words, intermediate were the trochaic accent I words, and the most difficult were accent II words.

Regarding incorrectly pronounced words, the result was that wrong stress placement produced larger loss, less 'yes' answers and more 'no' answers than wrong tonal word accent. When the 'yes' responses of mispronounced words were compared with correctly pronounced words, we saw a highly significant difference between correctly pronounced stress and mispronounced stress, for both types of stress change. There was a significant difference of amount of 'yes' answers between correct accent I and accent I as accent II, but not vice versa. This suggests that identification and comprehensibility of speech is more affected by wrong word stress placement than wrong word accent.

The study also shows that experimental methods combine well with phonetic, phonological and pedagogical issues. In further studies, we will test the perceptual weight of the third prosodic distinction of Swedish, quantity contrast, in relation to the two contrasts in the present study, and with respect to different positions in the word.

Pedagogical implication

Since the present experiment implies that the stress pattern of Swedish is more crucial for comprehensibility than tonal word accent, we suggest that second language learners of Swedish can benefit more from proficiency in perceiving and producing stress pattern. We can imagine a second language learner of Swedish going to school outside the Stockholm (capital) region. Her teacher may use a teaching material that describes the general Swedish stress patterns and also the Stockholm

variety of the word accent contrast. In addition to this, the teacher may unintentionally introduce her own local accent, despite her effort to comply with the tonal patterns described in the material. Even if the teacher succeeds to mimic the Stockholm tonal patterns, the learner will probably receive diverse input of tonal word accents from society outside school and from the media, as well. This may confuse her, not allowing her to discern what 'correct' Swedish word accent patterns are. The results of the present study suggest that the learner in the hypothetical situation, who is very likely to represent actual learners, can reduce her confusion and acquire appropriate pronunciations successfully, when the focus of teaching and learning lies on stress placement rather than on tonal word accents.

References

Abelin, Å., & Thorén, B. (2015). What affects recognition most – wrong word stress or wrong word accent? In *Proceedings from Fonetik 2015*, Institutionen för lingvistik, Lunds universitet.

Bannert, R. (1980). Svårigheter med svenskt uttal: Inventering och prioritering. In *Praktisk lingvistik nr 5.* Institutionen för lingvistik, Lunds Universitet.

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*. 2-3/1994

Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics, 23*(1), 83-103.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*, 71-102.

Munro, M. J., & Derwing, T. (1995). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning, 45*, 73-97.

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.