This book chapter is published in Babatspuli, E. & D. Ingram: *Phonology in Protolanguage and Interlanguage* (2018), Equinox. It is under copyright and the publisher should be contacted for permission to reuse or reprint the material in any form.

11
The perceptual weight of word stress, quantity and tonal word accent in Swedish

Åsa Abelin and Bosse Thorén

Introduction

Wherever migration or travelling takes place, people need to learn new languages. This learning entails a variety of interlanguages. Irrespective of whether you are a learner or a teacher of a language, you need to decide how to allocate time and effort into developing different sub-skills of the language. Four main skills are considered in second-language teaching and learning; listening, reading, speaking and writing. Proficiency in speaking requires sub-competences, such as pragmatic competence, fluency or making a clear pronunciation. Further, each of these sub-competences for speaking requires sub-skills. For example, to have a 'good' pronunciation, one needs to realize segmental features well: phonemes, phonotactics, assimilations, as well as prosodic features: rhythm and intonation.

In most cases, young children learning their first language (L1) as well as additional languages (L2s) acquire these pronunciation skills without formal training and often reach native-like speech also in additional languages. By contrast, adult learners of an additional language seldom reach native-likeness in their pronunciation of the language. However, ideally, they can still achieve a fluent, intelligible and well-received pronunciation of the language.

The present chapter is concerned with the pronunciation of Swedish as an additional language, and in particular three phonemic prosodic contrasts, namely word stress, quantity and tonal word accent. We attempt to find out, among these three prosodic contrasts, which is more crucial than the others for making one's speech intelligible. That is, if the second-language learner cannot acquire all of them perfectly, which of them should

be given more priority in learning and teaching Swedish pronunciation? We also want to study whether or not a pronunciation that lacks or mispronounces one of these contrasts can still be well understood.

Here, we need to clarify two things for our study. Firstly, we exclude the aspect of segmental properties of pronunciation, focusing on the effect of deconstructing Swedish prosody, and looking at the three phonemic parts and their respective relevance to intelligibility. We, however, do not assume or suggest that either prosody or the aspect of segmental properties is more important for intelligible speech. Secondly, we use the term 'perceptual weight' in our study to make it clear that we are only concerned with intelligibility of pronunciation, leaving out other aspects of foreign accent. The degree of native-like pronunciation is not addressed in the study, nor is the question of attitudes towards foreign accented speech. According to Munro and Derwing (1995) certain aspects of foreign accent can decrease intelligibility to some extent, but the degree of foreign accent in general does not seem to correlate with further decreased intelligibility. We therefore believe that certain features in a foreign accent may be crucial to intelligibility, while others are not particularly so. In fact, Bannert (1980) suggested some phonological features of Swedish as more crucial for intelligibility than others. Likewise, Thorén (2008) discussed the importance of prioritizing among the different Swedish prosodic contrasts and their respective acoustic correlates for pedagogical purposes.

In our study, we measure the degrees of importance of the three Swedish prosodic contrasts for intelligible speech. To do this, we use Swedish native speakers' perceptions, by means of letting them listen to natural pronunciations of words mixed with distorted pronunciations. As mentioned earlier, the general purpose of our paper is not to suggest native likeness as the norm for acquiring Swedish pronunciation. Instead, we seek to shed more light on the relative importance of different phonological and phonetic features in order to develop '*lingua franca* core features'[1] for Swedish. Swedish is not as international as English is, but in the last half century Sweden has developed into a multi-accented speech community where a huge variety of different linguistic groups dwell together. In addition to the varieties of native dialects and accents, the country has now people with diverse accents, such as Finnish, Arabic, Turkish, Persian, Vietnamese and Somali, to name a few. According to Parkvall (2016), there are between 150 and 200 languages used as first languages in Sweden today. Therefore, arguing for intelligibility rather than native likeness as the ideal pronunciation goal

1 The concept of lingual franca phonetic core features was suggested by Jenkins (2002) in the context of using English as an international language.

(Derwing and Munro 2015) and focusing on phonetic core features in pronunciation (Jenkins 2002) are very much relevant to teaching Swedish as an additional language. Accordingly, we are in the line of developing criteria for assessing different pronunciations of the Swedish language in terms of intelligibility rather than the degrees of accentedness.

The research questions are as follows:

(1) To what extent does the distortion of stress pattern affect the intelligibility of Swedish words when measured against the perception of native listeners?
(2) To what extent does the distortion of quantity category affect the intelligibility of Swedish words when measured against the perception of native listeners?
(3) To what extent does the distortion of tonal word accent category affect the intelligibility of Swedish words when measured against the perception of native listeners?
(4) Among the distortions of the three prosodic contrasts, which ranks first, second and third, in terms of negatively affecting the intelligibility of Swedish words?

Theoretical considerations

General considerations in speech perception

A general question concerning this study is which type of perception model is relevant for the interpretation of our results. There are two issues involved in perception: what phonetic cues and what units of perception are there? In accordance with Marslen-Wilson and Welsh (1978), we believe that speech input activates words in a cohort from 'left-to-right' and that there is interactive activation from the lexicon (cf. McClelland and Elman 1986; Marslen-Wilson 1987). Prosody and especially intonation have not been addressed much in psycholinguistic modelling, but there are some important example studies. For instance, Zhou and Marslen-Wilson (1994) provided a model of spoken word recognition for Mandarin that incorporates both segmental and tonal layers.

Soto-Faraco, Sebastian-Galles and Cutler (2001) studied whether suprasegmental information can facilitate lexical access in Spanish, and the relation between segmental mismatch and mismatch of lexical stress. Experiments using intentional mispronunciations of truncated words

showed that listeners process segmental and suprasegmental information in exactly the same way. Soto-Faraco et al. (2001) interpret the results to support a model of spoken-word recognition where the activation process is sensitive to all acoustic information relevant to the language's phonology.

In a study on Swedish tonal word accents, Abelin and Suomi (1997a) used a word-spotting design with two-syllable compounds having *accent I* or *accent II* where the first syllable of the compound was segmentally nonsense. The results of Abelin and Suomi (1997b) showed that the recognition of a morpheme in the second part of the compound is greatly facilitated (having shorter reaction times) when the first syllable is pronounced according to the *accent II* tone, as opposed to the *accent I* tone. Thus, the *accent II* tonal movement on a one-syllable nonsense word predicts that at least one more syllable follows. In a study on truncated *accent I* and *accent II* words with a slightly different purpose and design, Felder, Jönsson-Steiner, Eulitz and Lahiri (2009) concluded that both the perception of surface tonal contours and the identification of entire words are speeded up more by *accent I* than *accent II*. They also argued that *accent I* is governed by the underlying lexical structure with tonal specification.

Morphology, such as different inflectional forms, can affect processing as was shown by Söderström, Horne, Frid and Roll (2016), for example. They examined the perceptions 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. It was found that *accent II*-inducing suffixes preceded by an *accent I* tone were more difficult to process compared to *accent I*-inducing suffixes preceded by *accent II*. This is interpreted to mean that there is a stronger relation between suffixes and *accent II* as compared to *accent I*, which could imply that *accent II* can indeed be very important for identification and comprehension in certain contexts.

Grosjean and Gee (1987) argue that stressed syllables are used to initiate lexical search. Also, Cutler and van Donselaar (2001) showed that Dutch listeners can effectively use stress placement in the recognition of spoken words, and that mismatching stress placement reduced word activation. With the exception of Abelin and Thorén (2015), we don't know of any similar studies on manipulation of stress placement or quantity in Swedish.

We thus have assumed that suprasegmental information affects speech perception and have investigated how mismatching stress placement, tonal word accent or syllable quantity affects processing. Listeners do use lexical information to restore degraded speech. The question here has been which prosodic phonological category is the most sensible to mispronunciation, for L1 Swedish listeners.

The three prosodic contrasts of Swedish

Standard Swedish has three prosodic phonological contrasts: stress placement, quantity and tonal word accent. Swedish word stress is determined by prominence contrasts between syllables that are mainly signalled by syllable duration (Fant and Kruckenberg 1994). Quantity contrast largely depends on the durational relation between the vowel of a stressed syllable and the following consonant, resulting in the two categories: /V:C/ and /VC:/. Meanwhile, tonal word accent is mainly signalled by changes in the F0 contour and the timing of those changes in relation to the main stressed syllable of the word, for which related categories are: *accent I* (acute) and *accent II* (grave).

Stress contrast is traditionally regarded as a dynamic contrast, perceived as having varying prominence among syllables. 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 exist. A smaller number of trisyllabic minimal pairs, such as *Israel* [ˋiːsrael] 'the state of Israel' and *israel* [ɪsraˊeːl] 'Israeli citizen', are also identified. Stress differences also create semantic contrasts on the sentence level in verb phrases containing either one of the verbs plus an unstressed preposition, as in ˈ***häl****sa på N.N.* 'greet somebody' or a verb plus an stressed particle *hälsa* ˈ***på*** *N.N.* 'visit somebody'. Just like in English, the perceived prominence of stressed syllables in Swedish relies more on temporal (duration) and tonal (F0) acoustic correlates than dynamic (intensity) ones (Fant and Kruckenberg 1994; Fry 1958). In Swedish in particular, duration is the most reliable correlate (Fant and Kruckenberg 1994), while in English, it is change in F0 that is the most reliable acoustic cue to stress (Fry 1958). This entails that a syllable perceived as stressed is always longer than the same sequence of segments in unstressed position, and that higher sound intensity may or may not contribute to the impression of syllable prominence. Although Fant and Kruckenberg (1994) also conclude that F0 gestures, voice source parameters and differences in vowel quality combine with duration to signal syllable prominence, dynamic dimension can still be a possible cue to the listener's perception. Some studies in the extant literature, however, regard the Swedish stress contrast as a temporally based one. For example, Thorén (2008: 109) found that word stress had to be signalled by at least temporal properties in order to be correctly perceived by native listeners. In addition to the main temporal/durational acoustic correlates of the perceived prominence contrasts, stress is allocated, along the

timeline, to a word in a sequence of words and to a syllable in a sequence of syllables within the word.

In most regional varieties, quantity contrasts result in the two categories: /V:C/ as in *mäta* [mɛ:ta] 'to measure' and /VC:/ as in *mätta* [mɛt:a] 'measured (plural)'. Some studies report only /V:/ and /V/ in some South Swedish dialects (e.g. Gårding et al. 1974). In the latter case, the long vowel is mostly signalled by a diphthong in addition to pure duration. The quantity contrast is also accompanied by spectral differences between long and short vowel allophones. These differences are substantial for some vowel phonemes but very small for others. There is also a substantial variation between different regional varieties in this respect (cf. Behne, Czigler and Sullivan 1997; Bruce 2010; Hadding-Koch and Abramson 1964; Thorén 2003). Standard Finland-Swedish is known to have no or minimal spectral difference between long and short vowel allophones (Reuter 1971). There has been a debate regarding the Swedish quantity distinction, with some suggesting it is mainly a vowel distinction, some arguing it to be mainly a consonant distinction, and others saying it is equally based on vowel and consonant duration. It has also been suggested by Malmberg (1956), for example, that long and short vowels are separate phonemes, and that the mentioned durational contrasts in vowel and following consonants are there but are not phonemic per se. The present study, however, is not concerned with the refinement of theoretical accounts for the three contrasts; we merely aim to study their respective contribution to intelligibility.

A common proposal is that Swedish has two tonal word accent categories: *accent I* (acute), as in *tomten* [ˊtɔm:tən] 'the plot', and *accent II* (grave), as in *tomten* [ˋtɔm:tən] 'Santa Claus' (see Elert 1970), although only the grave accent can be considered as 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 di- or polysyllabic word), thus having a cohesive function, and it is limited to the word, simple or compound. Word accent is connected with a primary stressed syllable. When pronounced in isolation, words usually carry sentence accent, and *accent II* then tends to involve two F0 peaks (see Bruce 2010 for regional variation in tonal patterns for *accent I* and *II*). The two lower prominence levels in a sentence stress perspective – unstressed and secondary stressed – do not result in any signalling of word accent. The non-focal but still stressed 'accentuated' level (see Bruce 1977) usually results in a tonal fall within the stressed syllable in *accent II* words. In focal position, a tonal rise is added to a following syllable. The tonal contour of *accent I* seems to depend more on sentence intonation than on the word proper, although alternative views have also been advanced (e.g. Felder et al. 2009). Standard Finland

Swedish lacks the word accent contrast, which is also neutralized in singing. Irrespective of which word accent category should be seen as specified or unspecified, they appear with different F0 patterns in accentuated and focal positions. The reader may refer to Bruce (1977) for an extensive account of the Swedish tonal properties on sentence and word level.

Intelligibility aspects of Swedish prosody

As mentioned before, the present study does not promote native-likeness for determining whether the prosodic features of a pronunciation are correctly or incorrectly signalled. It is only concerned with how much they affect intelligibility. This is in line with Derwing and Munro (2015), Abercrombie (1949) and others, who propagate a shift in the perspective from native-likeness to intelligibility as the goal of second-language learning, particularly for adult learners. Furthermore, as suggested by Jenkins (2002), we assume that some phonetic features of a language are more crucial for intelligibility than others. This idea is not completely new in the context of Swedish as an additional language. For example, Bannert (1980) suggested word stress, sentence stress, 'vowel quantity' and increased duration of stressed syllables as important goals in pronunciation teaching and learning. He also proposed that tonal word accent, complementary consonant length and some assimilations could be given less priority. However, his proposal was mostly based on his own and colleagues' intuition, lacking robust empirical evidence. Recently, Abelin and Thorén (2015) empirically confirmed Bannert's ranking concerning word stress and tonal word accent. In the present study, where investigating quantity contrast is also added, we do not separate 'vowel quantity' from the complementary consonant length, but we regard the quantity distinction as an entity involving the elements mentioned above. In future studies, we hope to look at the perceptual weight of segmental features, such as voicing, consonant clusters and some assimilations, and their respective contributions to intelligibility.

There is some structural evidence that word stress and quantity should play an more important role in the perception and comprehension of Swedish than tonal word accent. As mentioned above, the two former contrasts are present in all varieties of Swedish, which is not the case for tonal word accent. There are some jingles and joking sentences in Swedish where word stress is changed, resulting in total loss of intelligibility to the first-time listener. Gårding (1979: 13) describes what is assumed to happen when the native Swedish listener is confronted with a word that has distorted

stress when produced by an L2 speaker (Translation from Swedish by the present authors):

> How about [ɛəlá:re] (instead of [ɛɛ̀l:ʁrə] 'cellar')? The word loses its identity. The listener searches for a similar word, i.e. a word with the same stress pattern, rummages around in the brain-lexicon, but finds no correspondence/equivalence. As you see [Gårding refers to a table] it is supposed to mean *källare* 'cellar'.

The prosodic features of Swedish, however, can be associated with the morphological structure of the words. As shown by Söderström, Roll and Horne (2012), the tonal word accent pattern of the first syllable of a word as well as stress placement may entail morphological information that makes a word easier or more difficult to process.

In our study we have not tested the effect of morphology, or the effect of other contexts such as syntactic, semantic or pragmatic.

We currently look at intelligibility from a native Swedish listener's perspective, but in future studies we want to look at intelligibility in all combinations of speakers and listeners, involving L1 speakers to L2 listeners, L2 speakers to L1 listeners, and L2 speakers to L2 listeners. L2 listeners and speakers should be understood as language users from diverse L1 backgrounds.

Although using an L1 speaker of Swedish for the production of distorted prosodic categories cannot mimic exactly how L2 speakers typically mispronounce the current prosodic categories, it still allows a higher degree of control with respect to the phonetic properties that are studied than using an L2 speaker.

Pedagogical considerations for Swedish L2 pronunciation

According to Gårding (1974) the prosody of an additional language is something that is particularly hard to be learned by the adult learner. A general account of second-language pronunciation learning at different ages is provided by Derwing and Munro (2015). They show that it is possible for adults to perceive and produce new segmental and prosodic features. Our study does not address how these adult learners should best be helped to achieve the prosody in an L2, but we want to find out whether and how the three Swedish prosodic contrasts differ with respect to their importance in making speech intelligible.

As Kjellin (1995) and Thorén (2008) suggest, the properties of stress and quantity can be combined to benefit teaching pronunciation of Swedish as a second language. Since both contrasts rely mainly on duration as per-ceptual cues, exaggeration of the duration of the stressed syllables, which thereby signals the prominence of syllables, can be utilized in teaching pro-nunciation. Ffurthermore, if lengthening of the stressed is realized by lengthening the correct segment, (vowel or following consonant), learners are assumed to have learned the realization of the intended quan-tity category. We can talk about killing two birds with one stone, i.e. we promote two prosodic contrasts by means of one measure – lengthening the correct segment in the correct syllable. Our previous study showed that stress is important for intelligibility, and if the present study shows that the quantity distinction is also important, teachers and learners can confidently combine the two prosodic contrasts by means of their shared acoustic cue, i.e. duration.

Outline of the study

In this study, we have carried out three experiments that involve all three Swedish phonemic prosody contrasts – stress, quantity and tonal word accent. The aim was to determine a ranking order for the three Swedish prosodic contrasts with respect to their relative importance for intelligibil-ity. We prepared a number of disyllabic words: some are real words, some are nonsense words, and some are real words pronounced in the opposite phonological category, i.e. (1) changing the word stress category from tro-chaic to iambic and vice versa, (2) changing the quantity category from /V:C/ to /VC:/ and (3) changing the tonal word accent category from *accent I* to *accent II* and vice versa. We avoided minimal pairs to prevent creat-ing another known word by changing category.[2] We are aware that this manipulation is somewhat crude and does not reflect all typical L2 reali-zations of stress patterns. For example, whereas we changed words to the opposite phonological category with typical Swedish clear realization of acoustic correlates, some L2 speakers may also typically produce unclear realizations, resulting in ambiguous prosodic categories, presumably due to insufficient awareness of the mentioned prosodic patterns. The method used in the study consists of three lexical decision experiments (1a, 1b and

2 Please see the discussion of Experiment 1, which acknowledges minor mistakes in this process

2) with word stimuli in both the correct and opposite prosodic category. We measured degree of correct identifications, rejections and reaction times.

Experiments 1a and 1b

Material and design

The material for Experiment 1a consisted of three sets of intact words: 10 trochaic *accent I* words, e.g. bilen [ˊbiːlen] 'the car', 10 trochaic *accent II* words, e.g. gatan [ˋgɑːtan] 'the street', and 10 iambic words, e.g. kalas [kaˊlɑːs] 'party'; two sets of words with changed prosodic category: 10 originally trochaic words pronounced with iambic stress, e.g. låset* [loˈsɛːt] 'the lock', and 10 originally *accent II* words pronounced with tro-chaic stress and *accent I*, e.g. sagan* [ˊsɑːgan] 'the fairy tale'. As distractors, we presented 26 disyllabic non-words, with varying stress patterns or tonal accent categories. Likewise, the material for Experiment 1b consisted of three sets of intact words: 10 trochaic *accent I* words, e.g. köket [ˊɕøːkət] 'the kitchen', 10 trochaic *accent II* words, e.g. gatan [ˋgɑːtan] 'the street', 10 iambic words, e.g. kalas [kaˊlɑːs] 'party', 10 originally iambic words pronounced with trochaic stress, e.g. kanel *[ˊkanel] 'cinnamon', and 10 originally *accent I* words pronounced with *accent II*, e.g. degen *[ˋdeːgən] 'the dough'. The same 26 disyllabic non-words were used in both parts of this experiment. Thus, words with *accent I* were mispronounced with *accent II* and vice versa, and words with trochaic stress were mispronounced with iambic stress and vice versa. See Appendix I for the complete list of stimulus words.

All the trochaic words (with the exception of sälar, 'seals') were nouns in the definite singular form, having excluded possible members of minimal phonological pairs.[3] The words were recorded by a male phonetician with a moderate Stockholm dialect. This means that his pronunciation cannot be traced with certainty to Stockholm but rather to a wider area in eastern Sweden. Recordings were made with a Røde NT3 condenser microphone to a laptop at a sampling frequency of 22,050 Hz in a silent studio at the University of Umeå, Sweden, and editing was made with Praat (Boersma and Weenink 2013).

3 We accidently included one word *[ˋjuːrət], whose distorted form can be interpreted as a real word. This is described in detail in the general discussion.

There was some deliberation about how to treat vowel quality in the stressed and unstressed syllables, since these vary according to the degrees of stress and quantity category. We decided to choose vowels which do not vary so much in unstressed vs. stressed position or short vs. long allophones (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) as much as possible. Each word was presented until self-terminated, in all cases just below 1,000 ms. At the same time, participants had 1,000 ms to react to each stimulus. Thus, the time allotted for reaction to the stimuli started when the word started and there was a 1,000 ms pause between each word. For building and running the experiment, the PsyScope software was used (Cohen, MacWhinney, Flatt and Provost 1993).

Procedure and participants

Two lexical decision tests were performed in Experiments 1a and 1b. 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 1a, one by one in random order. In Experiment 1b, 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 1b, one by one in random order. The subjects were instructed to press one key on the 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 1,000 ms period were categorized as loss. The subjects had no reported hearing impairment.

Results

Accuracy

Figure 11.1 shows the main results of Experiments 1a and 1b. It turned out that the task was quite demanding, and that the loss in the experiment was large.

It is evident from Figure 11.1 that wrong stress placement produced more rejections than wrong tonal word accent in both Experiments 1a and 1b. 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 ($p < 0.0001$). The difference in the number

of 'yes' responses between *accent I* mispronounced as *accent I* and *accent II* mispronounced as *accent I* was not significant. There was no significant difference between the responses for trochaic pronounced as iambic and iambic pronounced as trochaic.

Figure 11.2 shows a comparison of the wrongly pronounced words with the correctly pronounced words. The correctly pronounced words are, as expected, the most robust; they exhibit 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 the 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 ($p = 0.0233$). The difference in the number of 'yes' responses between correctly pronounced *accent I* 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 more loss than the reverse condition.

The difference in the number of 'yes' responses between correctly pronounced trochaic words and trochaic words pronounced with iambic stress was significant ($p < 0.0001$). Likewise, the difference in the number of 'yes' responses between correctly pronounced iambic words and iambic words pronounced with trochaic stress was significant ($p < 0.0001$).

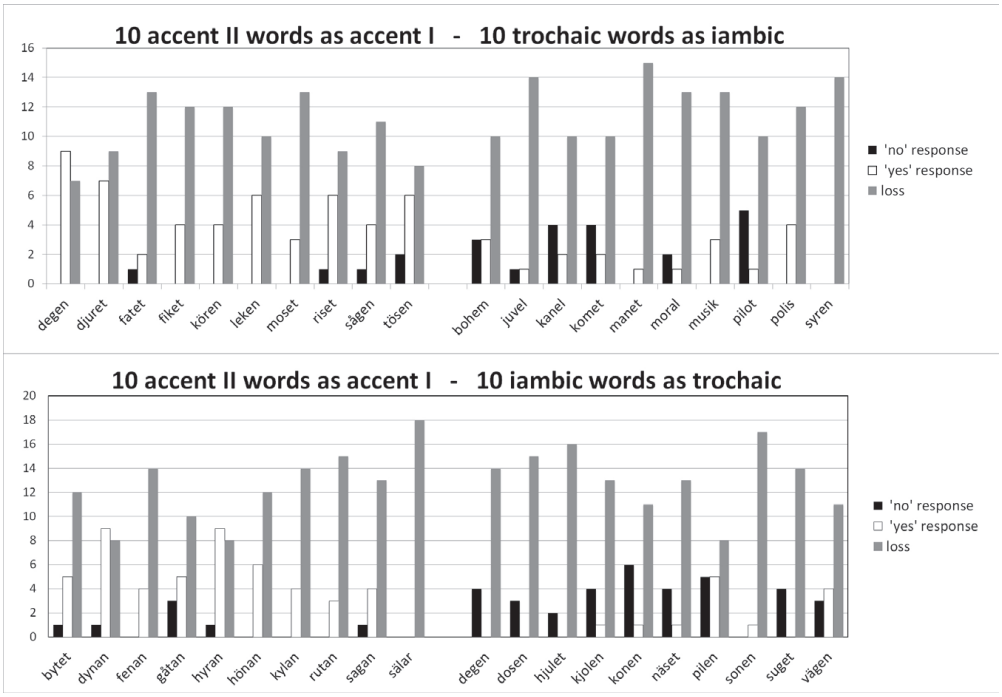

Figure 11.1. Results of Experiment 1.

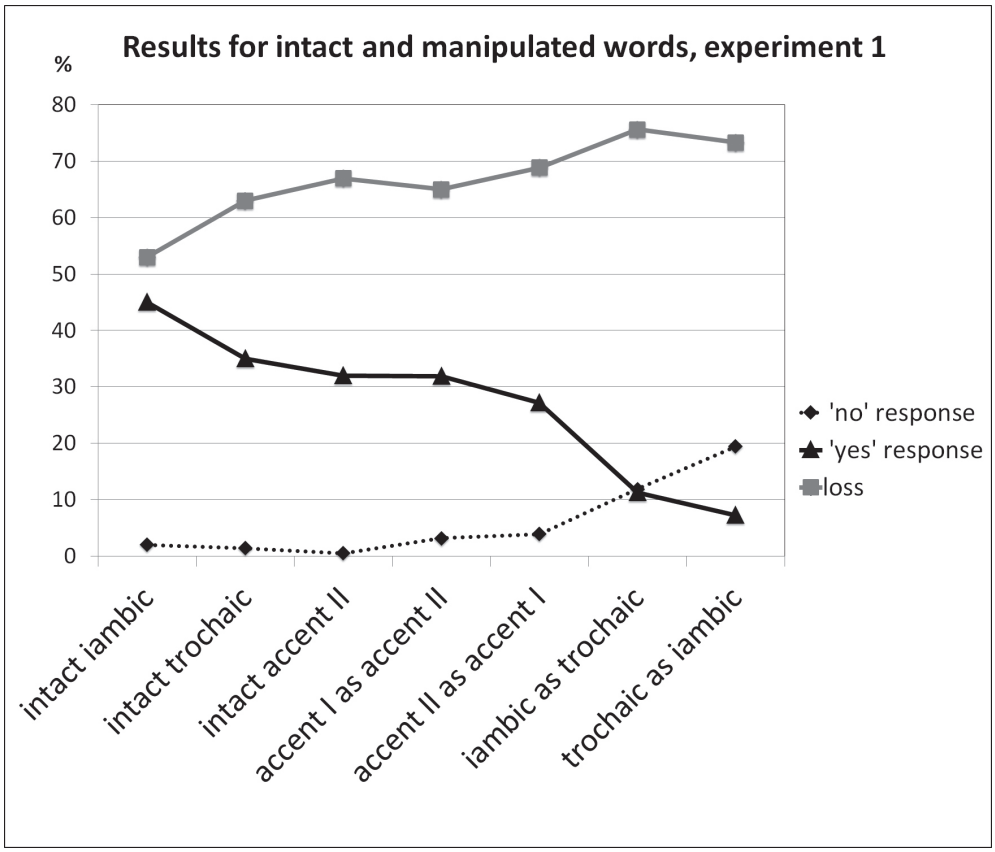

Figure 11.2. Comparison of wrongly and correctly pronounced words: Experiment 1.

There was a connection between loss and 'no'/'yes' responses. There was a negative correlation between number of 'yes' responses and loss (r^2 = 0.8473). Furthermore, the loss was greater where there were more 'no' responses.

Reaction times

It was not possible to compare reaction times in 'yes' responses' for word accent errors and stress placement errors, 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 word durations, but not if 200 ms were deducted for motor activation. There was a tendency for less

loss, and the 'yes' responses were more numerous when the durations were shorter.

Discussion

The results of Experiment 1 suggest greater perceptual weight in the case of the stress pattern when compared to tonal word accent.

There was a connection between loss and 'yes' responses, with negative correlation between the number of 'yes' responses and loss. Furthermore, the loss was greater where there were more 'no' responses. 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 1,000 ms. But the result could also be due to an inability to interpret the wrongly pronounced word. This is further explored in Experiment 2, which allowed for longer reaction times.

Although we checked the words for membership in minimal phonological pairs, one such case became apparent after the first presentation of our results. The word *djuret* [ˊjʉːrət] 'the animal', normally pronounced with *accent I* and here pronounced with *accent II* *[ˋjʉːrət], can actually be interpreted as the compound word *djurrätt* [ˋjʉːˌrɛtː] 'animal rights'. As shown in Figure 11.1, the mispronounced ˋ*djuret* showed the second highest number of 'yes' answers and relatively less 'loss' in this distortion category, indicating that it was recognized as a real word by just above half of the listeners. We do not, however, know whether listeners perceived the word as 'the animal' or 'animal rights'.

Conclusion

Participants identified correctly pronounced words as real words more easily and they produced smaller loss, more 'yes' responses and less 'no' responses than they did with the wrongly pronounced words. In terms of the degree of ease among the correctly pronounced words, the iambic category was the highest, the trochaic *accent I* category was next, and the *accent II* category was the lowest.

Among the incorrectly pronounced words, wrong stress placement produced larger loss, fewer 'yes' answers and more 'no' answers than wrong tonal word accent. When the 'yes' responses of mispronounced words were compared with the 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 in the amount of 'yes' answers between correct *accent I* and *accent I* as *accent II*, but not vice versa. This suggests that intelligibility of speech is more affected by wrong word stress placement than by wrong word accent.

In Experiment 2, we tested the perceptual weight of the third prosodic distinction of Swedish, quantity contrast, together with the two contrasts in Experiment 1.

Experiment 2

Material and design

A lexical decision experiment was performed, where 10 native Swedish listeners were exposed to 50 intact words representing combinations of trochaic, iambic, *accent I*, *accent II* as well as /V:C/ and /VC:/ categories. The test words were 10 originally trochaic words pronounced with iambic stress patterns, 10 originally *accent I* words pronounced with *accent II*[4] and 10 trochaic /V:C/ words pronounced as /VC:/. It was not possible to include all three contrasts in both directions since the experiment would become too large. 60 nonsense words with the same combinations of phonological categories served as distractors. See Appendix II for the complete list of stimulus words.

The words were recorded by the same person as in Experiment 1 and under the same technical conditions, except for the place of recording. This time the recording was made in a small room with provisional anti-echo treatment. Each word was presented until self-terminated, and all words lasted just below 1,000 ms. Participants had 1,500 ms to react to each stimulus (500 ms longer than in Experiment 1). Reaction times were measured beginning the moment the words started. Between each word, there was a 1,000 ms pause. For building and running the experiment the PsyScope software was used (Cohen et al. 1993). Participants were instructed to judge as quickly as possible whether the words they heard were real words or not. They were asked to press the 'yes' button if they heard a real word, and the 'no' button if they heard a non-word. The number of yes/no answers

4 As we had received comments on the first experiment that *accent II* words pronounced with *accent I* could be perceived as 'correct' pronunciations in some dialects, we chose to make the distortion from *accent I* to *accent II*.

and non-responses (answers that exceeded the reaction time limit) were counted and reaction times were measured.

Results

The results of Experiment 2 are shown in Figure 11.3. As in Experiment 1 participants identified correctly pronounced words as real words more easily and they produced more 'yes' responses, fewer 'no' responses and less loss than they did in the wrongly pronounced words. The results also show that participants tended to judge words as non-real to a higher degree when pronounced with distorted quantity than when pronounced with distorted stress and distorted tonal word accent. The frequency of non-responses (loss) and non-word decisions for distorted word stress was slightly lower than for distorted quantity but still much higher than for distorted word accent.

Through pairwise analysis of correct stress placement with distorted stress placement, correct quantity with distorted quantity, and correct

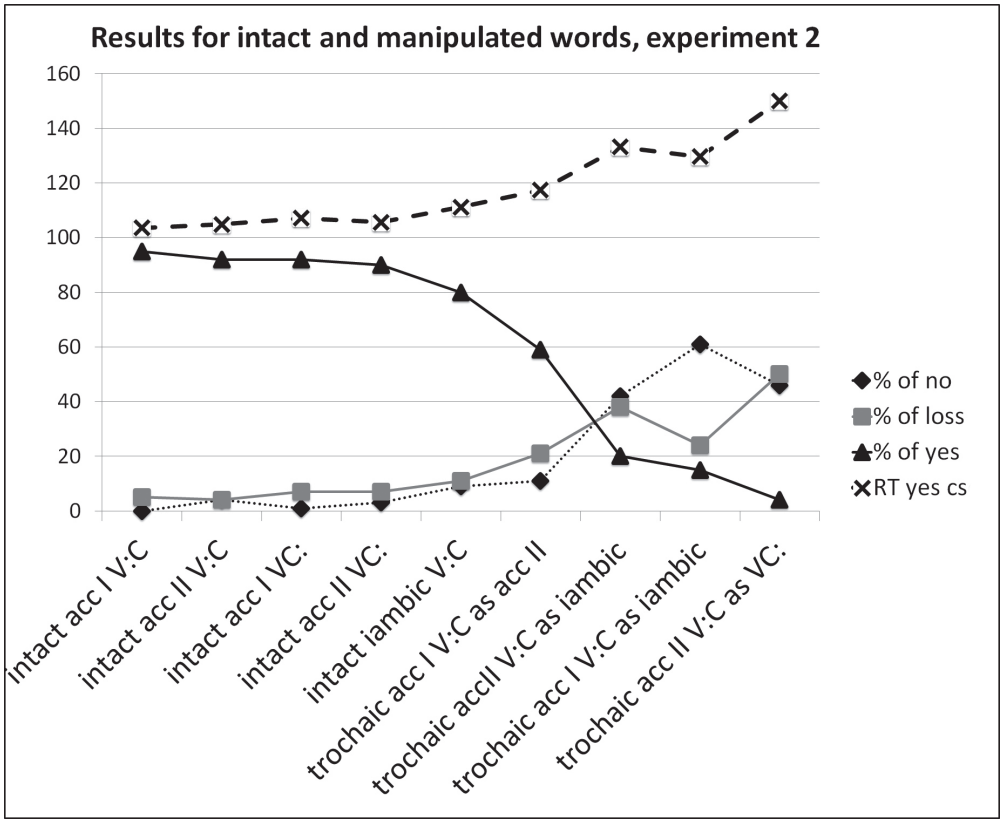

Figure 11.3. Comparison of wrongly and correctly pronounced words: Experiment 2.

tonal word accent with distorted word accent, we see that distorted quantity had the largest negative effect on word identification. There was a significant difference in the amount of 'yes' responses between correct and incorrect quantity (unpaired *t*-test, $p = 0.0001$). The difference in reaction times could not be calculated for wrongly pronounced quantity since so many subjects did not react within the allotted 1,500 ms. (The mean reaction time [RT] of 150 centiseconds for wrong quantity in the diagram is really a dummy, since reaction times were much longer and, thus, not measurable.) These long reaction times indicate great difficulties in processing. The difference in 'yes' and 'no' answers for trochaic pronounced incorrectly with iambic stress was also significant (unpaired *t*-test, $p = 0.0001$). The difference in 'yes' and 'no' answers for incorrect *accent I* as *accent II* was significant as well (unpaired *t*-test, $p = 0.002$), as was the difference in reaction time for correct trochaic words and trochaic words pronounced incorrectly with iambic stress (unpaired *t*-test, $p = 0.0001$). The difference in reaction time for incorrect *accent I* to *accent II* was also significant (unpaired *t*-test, $p = 0.0085$), though to a lesser degree as expected.

We can also see that reaction times show a negative correlation with 'yes' judgements (–0.967) and a positive correlation with 'no' judgements (0.875). In other words, those distorted words which produced more 'no' responses also had longer reaction times when they were judged as real words. Both 'no' responses and longer reaction times indicate difficulty in identification and, therefore, lower intelligibility of the mispronounced words.

The results of Experiment 2 show that loss is diminished when longer reaction times are allowed. Reaction times for 'yes' responses became possible to measure and the length of reaction times largely reflects the same order as the amount of 'yes' and 'no' answers.

Discussion and conclusion

The results of Experiment 2 indicate that distorted quantity is more detrimental to word identification than distortions of both word stress and tonal word accent. However, both word stress and quantity place themselves near each other and with some distance from word accent contrast with respect to listeners' sensitivity to changed category. Thus, the results from Experiment 1 were replicated, adding the results on quantity contrast.

General discussion

The results of this study suggest greater perceptual weight for quantity and stress pattern when compared with tonal word accent. Change in the quantity category was suspected to be less detrimental to word recognition than change in the stress pattern, but quantity turned out to be the most crucial of the three tested prosodic contrasts.

There was a negative correlation between the number of 'yes' and 'no' responses in both experiments. Furthermore, there was a negative correlation between reaction times and 'yes' responses in the second experiment. Altogether, these findings indicate that recognition of words with mispronounced prosody is difficult. We also saw that quantity errors were the most harmful to recognition and that tonal word accent errors were the least harmful, although not negligible.

These results could also be discussed in relation to the left-to-right cohort model of speech perception (cf. Marslen-Wilson 1987), but it is unclear how prosody can get accommodated in this model. One question is whether an early absence of stress placement would be more detrimental for recognition than a late absence. That is, would a trochaic word with changed stress-placement (which ought to have stress on the first syllable – and therefore lacking early durational and intonational cues) be more difficult to process than an iambic word with changed stress-placement (which ought to have stress on the second syllable)? There is some evidence of this in Experiment 1, but more studies on Swedish are needed for addressing this question. Similarly for *accent I* and *accent II*, there were slightly more errors for *accent II* as *accent I* than for *accent I* as *accent II*. If *accent I* is characterized as later peak and *accent II* is characterized as earlier peak in isolated pronunciation (cf. Felder et al. 2009 on Stockholm dialect) this indicates that an absence of correct tone movement on the first syllable in a disyllabic word is more disturbing than the presence of wrong tone movement on the first syllable in a disyllabic word. An alternative explanation is that it is the absence of high tone which causes more problems in perception. This indicates that the absence of a cue (for stress or tonal word accent) is more disturbing to the listener than the wrong presence of a cue (for stress or tonal word accent), but this needs to be further verified.

We can only speculate whether a change in the quantity category in either direction between /V:C/ and /VC:/ would have a more or less detrimental effect on word recognition, since we only included the direction from /V:C/ to /VC:/ in our study. We assume, however, first, that identification of the quantity category is processed from 'left-to-right', since vowel

quality must be the first perceptual cue to the listener. Vowel duration and post-vocalic consonant duration are subsequently assumed to confirm or falsify the listener's original hypothesis. Establishing the quantity category is a first step, but word identification still remains to be done. Again, we can only speculate, but we assume that word identification is dependent on the number of known homophonic first syllables that are parts of real words. On the other hand, in a /VC:/ word, the post-vocalic consonant is revealed within the first syllable giving the listener one more cue to the entire word (or non-word). In the time frame within which this is done, the /V:C/ word reveals only the vowel. That could possibly result in intact /VC:/ words being easier to process than /V:C/ words. Our data, however, do not show any significant differences in this respect. But when we change either of the two categories into the other, and focus on word recognition rather than perception of the quantity category, the listener is likely to depend on the frequency of different stems that may agree with the perceived sequences.

The words of the present experiments were not checked for frequency or number of phonological neighbours. The reason they 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. Söderström, Horne and Roll (2016) found that differences in processing of *accent I* and *accent II* stems can in part be explained by the density of the phonological neighbourhood of stems, due to lexical competition. In the present experiment, we cannot draw similar conclusions regarding lexical effects, since we do not have reliable data on phonological neighbourhoods.

What is puzzling is that, in our first experiment, 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 is in opposition to the result in Experiment 2. A post-check of frequencies for the stimuli words in the two tests did not provide any explanation for this.

In relation to studies concerning the effect of morphology, such as that of Söderström et al. (2012), a question arises whether *accent II* might be more important to comprehension when there are other errors as well, such as in the speech of learners of Swedish as a second language who may use the wrong suffixes in nouns or verbs. Adding further learner errors, such as word order mistakes or wrong lexical choices, complicates the picture further.

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

judged out of context. Furthermore, as mentioned above, L2 speakers do not typically change any of the mentioned prosodic categories into a clear realization of the opposite category, but rather into a generally unclear pronunciation with respect to the mentioned prosodic contrasts. We are therefore planning follow-up studies in more natural scenarios with combinations of L1 and L2 speakers and listeners.

Conclusion

We conclude that Swedish L1 listeners perceive and identify words with incorrect quantity category, incorrect stress placement and incorrect tonal word accent with greater difficulty than words pronounced with correct stress, quantity and word accent. Thus, correctly pronounced words are easier to identify (they produce smaller loss, more 'yes' responses and less 'no' responses, and entail shorter reaction times) than the wrongly pronounced words.

For the incorrectly pronounced words the results show that wrong quantity category and wrong stress placement produce fewer 'yes' answers, more 'no' answers and more loss than wrong tonal word accent. The study also shows that phonetic, phonological and psycholinguistic experimental methods combine well for dealing with pedagogical issues.

Pedagogical implication

Based on the results, we suggest that learners of Swedish as a second language benefit more from proficiency in temporal prosodic properties than in the choice of word accent category or precise realization of word accent category (cf. Thorén 2008). In fact, word accent categories are realized differently in different geographical regions, and some varieties do not utilize the contrast at all.

Since the second experiment implies that quantity and stress pattern in Swedish are more crucial to intelligibility than tonal word accent, we suggest that second-language learners of Swedish are specifically trained in perceiving and producing both the quantity distinction and the stress pattern. We can imagine a second-language learner of Swedish going to school outside the Stockholm (capital) region. Her teacher may use teaching material that describes the general Swedish stress and quantity 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 in mimicking the Stockholm tonal patterns, the learner will probably receive diverse input on tonal word accents from social interaction outside the school and from the media, as well. This may confuse her interlanguage system, not allowing her to discern what the 'correct' Swedish word accent patterns are. The results of the present study suggest that the learner in this hypothetical situation, who is very likely to represent actual learners, can minimize confusion and successfully acquire appropriate pronunciation, if the focus of teaching and learning is placed on the temporal prosodic properties rather than on tonal word accents.

In addition to a better foundation on what should be prioritized in teaching Swedish L2 pronunciation, teachers and learners of Swedish as a second language can draw on the findings of Fant and Kruckenberg (1994) and the suggestions by Kjellin (1995) and Thorén (2003, 2008) to lengthen the correct segment (vowel or consonant) in the stressed syllable, thereby promoting the significance in the signal of both stress placement and the quantity category of a given word. Teaching and learning can benefit from the finding that two important phonological contrasts share the same main acoustic cue, namely, duration.

Further experiments

We believe that future studies should examine the relative perceptual weights of segmental features like vowel quality, consonant features, phonotactic features and the role of assimilations and reductions; also intelligibility tests should involve phrases and sentences in addition to single words. Together with studies of L1 and L2 speakers and listeners, we intend to replicate the two experiments in this study with typically developing Swedish children aged around 4 to 6 years. We will then be able to see how children's word recognition is affected by non-standard prosody. The question is whether speech perception in children is more or less segmental than speech perception in adults, thus, addressing the effects in protolanguage. A study by Sundström, Samuelsson and Lyxell (2014) on word repetition and non-word repetition alluded to age-dependent differences regarding how prosody is stored and integrated with segments.

Appendix I

The categories of real/intact words, nonsense words and words with changed prosodic categories that are used in Experiment 1. The manipulated categories are shown in bold.

Intact real words			*Nonsense words*		
10 intact words trochaic *accent II* /V:C/	10 intact words trochaic *accent I* /V:C/	10 intact words iambic *accent I* /V:C/	26 Nonsense /V:C/, mixed *accent I*, *accent II*, trochaic and iambic		
bulan	bilen	baron	`blyran	`göpan	púret
bönan	boken	belag	búget	kadél	tö´ket
diket	bordet	besök	dában	`kafan	sirán
duvan	duken	cerat	`dyset	`kogan	vakós
dåren	dalen	dekal	fakén	kýgen	`tjoman
faran	fåret	filur	femól	linár	topít
gatan	fölet	metan	fúket	míben	
kakan	filen	minut	gáket	misýt	
leran	gåsen	raket	`garan	núgen	
ligan	huset	safir	golát	porít	

Changed phonological category			
10 **trochaic** *accent I* /V:C/ pron. as **iambic** *accent I* /V:C/	10 **iambic** *accent I* /V:C/ pron. as **trochaic** *accent I* /V:C/	10 **trochaic** *accent I* /V:C/ pron. as **trochaic** *accent II* /V:C/	10 **trochaic** *accent II* /V:C/ pron. as **trochaic** *accent II* /VC:/
degen	bohem	degen	bytet
dosen	juvel	djuret	dynan
hjulet	kanel	fatet	fenan
kjolen	komet	fiket	gåtan
konen	manet	kören	hyran
näset	moral	leken	hönan
pilen	musik	moset	kylan
sonen	pilot	riset	rutan
suget	polis	sågen	sagan
vägen	syren	tösen	sälar

Appendix II

The categories of real/intact words, nonsense words and words with changed prosodic category that are used in Experiment 2. The manipulated categories are shown in bold.

Intact real words					*Nonsense words*		
10 intact words trochaic *accent II* /V:C/	10 intact words trochaic *accent II* /VC:/	10 intact words trochaic *accent I* /V:C/	10 intact words trochaic *accent I* /VC:/	10 intact words iambic *accent I* /V:C/	20 non-sense words trochaic *accent I* /V:C/+ /VC:/	20 non-sense words trochaic *accent II* /V:C/+ /VC:/	20 non-sense words iambic *accent I* /V:C/
bulan	himmel	bilen	buggen	baron	buget	byran	bynet
bönan	hackan	boken	bussen	belag	dyset	garan	diran
diket	kaffe	bordet	dammen	besök	föket	höpar	dugåm
duvan	killen	duken	luggen	cerat	gaket	kalan	faken
dåren	soffan	dalen	lasset	dekal	kupet	nögat	femål
faran	villan	fåret	sucken	filur	tjygen	pagar	filås
gatan	mössan	fölet	missen	metan	lafen	tjipan	golöt
kakan	ärtan	filen	lacken	minut	miset	jöpat	henut
leran	bullen	gåsen	tuppen	raket	nugen	kafan	kadel
ligan	pannan	huset	lappen	safir	töben	säpan	linar
					sibben	möckan	matus
					byllet	våmmar	misyt
					dacket	faggar	mokut
					fippet	koggan	piret
					gåppen	paffan	porit
					marret	vuggan	potil
					nyppen	dibban	siben
					påffet	gåppan	siran
					rybben	jöllan	topit
					vellet	pyffan	vakos

Changed phonological category			
10 **trochaic** *accent I* /V:C/ pron. as **trochaic** *accent II* /V:C/	6 **trochaic** *accent II* /V:C/ pron. as **iambic** *accent I* /V:C/	10 **trochaic** *accent I* /V:C/ pron. as **iambic** *accent I* /V:C/	10 **trochaic** *accent II* /V:C/ pron. as **trochaic** *accent II* /VC:/
basen	logen	djuret	bön:an
bogen	medar	dosen	näs:an
degen	sagor	fyren	sön:er
diset	sätet	kilot	fån:en
fatet	sidor	kjolen	kål:or
gamen	kilar	låset	ler:an
sylen		näset	nit:ar
låret		polen	skid:an
piken		renen	tjäl:en
leken		rågen	våg:ar

References

Abelin, Å. and Suomi, K., 1997a, 'Swedish word accents and speech segmentation', in *Proceedings of FONETIK 97*, Umeå University.

Abelin, Å. and Suomi, K., 1997b, 'Prediction of Swedish tonal word accent', Department of Linguistics, University of Gothenburg (unpublished manuscript).

Abelin, Å. and Thorén, B., 2015, 'The relative weight of two Swedish prosodic contrasts', in E. Babatsouli and D. Ingram (eds.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech*, pp. 3–7, Chania, Greece (ISBN: 978-618-82351-0-6).

Abercrombie, D., 1949, 'Teaching pronunciation', *ELT Journal* 3. https://doi.org/10.1093/elt/iii.5.113

Bannert, R., 1980, 'Svårigheter med svenskt uttal: Inventering och prioritering', in *Praktisk lingvistik nr 5*, Institutionen för lingvistik, Lunds universitet.

Behne, D.M., Czigler, P.E. and Sullivan, K.P.H., 1997, 'Swedish quantity and quality: A traditional issue revisited', in *Phonum 4*, Department of Linguistics, Umeå University.

Boersma, P. and Weenink, D., 2013, *Praat: Doing phonetics by computer*, retrieved from http://www.praat.org.

Bruce, G., 1977, *Swedish word accents in sentence perspective*, Vol. 12, Lund University.

Bruce, G., 2010, *Vår fonetiska geografi – Om svenskans accenter, melodi och uttal*, Lund: Studentlitteratur.

Bruce, G., 2012, *Allmän och svensk prosodi*, Lund: Studentlitteratur.

Cohen, J.D., MacWhinney, B., Flatt, M. and Provost, J., 1993, 'PsyScope: A new graphic interactive environment for designing psychology experiments', *Behavioral Research Methods, Instruments, and Computers* 25(2), 257–71. https://doi.org/10.3758/BF03204507

Cutler, A. and van Donselaar, W., 2001, 'Voornaam is not (really) a homophone: Lexical prosody and lexical access in Dutch', *Language and Speech* 44(2), 171–95. https://doi.org/10.1177/00238309010440020301

Derwing, T. and Munro, M.J., 2015, *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*, Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/lllt.42

Elert, C.-C., 1970, *Ljud och ord i svenskan*, Almqvist & Wiksell: Stockholm.

Fant, G. and Kruckenberg, A., 1994, 'Notes on stress and word accent in Swedish', *STL-QPSR* 2–3.

Felder, V., Jönsson-Steiner, E., Eulitz, C. and Lahiri, A., 2009, 'Asymmetrical processing of lexical tonal contrast in Swedish', *Attention, Perception, and Psychophysics* 71(8), 1890–99. https://doi.org/10.3758/APP.71.8.1890

Fry, D.B., 1958, 'Experiments in the perception of stress', *Language and Speech* 1, 126–52.

Gårding, E., 1974, 'Den efterhängsna prosodin', in U. Teleman and T. Hultman (eds.), *Språket i bruk*, Lund: Liber.

Gårding, E., 1979, 'Avvikande uttal – Analys och värdering', in *Praktisk Lingvistik 1*, Institutionen för lingvistik, Lunds universitet.

Gårding, E., Bannert, R., Bredvad-Jensen, A-C., Bruce, G. and Naucler, K., 1974, 'Talar skåningarna svenska?', *Svenskans beskrivning* 8, 107–17.

Grosjean, F. and Gee, J.P., 1987, 'Prosodic structure and spoken word recognition', *Cognition* 25, 135–55. https://doi.org/10.1016/0010-0277(87)90007-2

Hadding-Koch, K. and Abramson, A., 1964, 'Duration versus spectrum in Swedish vowels: Some perceptual experiments', *Studia Linguistica* 18, 94–107.

Jenkins, J., 2002, 'A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language', *Applied Linguistics* 23(1), 83–103. https://doi.org/10.1093/applin/23.1.83

Kjellin, O., 1995, *Svensk prosodi i praktiken – instruktioner och övningar i svenskt uttal, speciellt språkmelodin*, Stockholm: Hallgren & Fallgren studieförlag.

Malmberg, B., 1956, 'Distinctive features of Swedish vowels: Some instrumental and structural data', in H. McLean et al. (eds.), *For Roman Jakobson*, pp. 316–21, The Hague: Mouton.

Marslen-Wilson WD., 1987, 'Functional parallelism in spoken word recognition', *Cognition* 25, 71–102. https://doi.org/10.1016/0010-0277(87)90005-9

Marslen-Wilson, W.D. and Welsh, A., 1978, 'Processing interactions and lexical access during word recognition in continuous speech', *Cognitive Psychology* 10, 29–63. https://doi.org/10.1016/0010-0285(78)90018-X

McClelland, J.L., and Elman, J.L., 1986, 'The TRACE model of speech perception', *Cognitive Psychology*, 18(1), 1–86. https://doi.org/10.1016/0010-0285(86)90015-0

Munro, M.J. and Derwing, T., 1995, 'Foreign accent, comprehensibility and intelligibility in the speech of second language learners', *Language Learning* 45, 73–97. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Parkvall, M., 2016, *Sveriges språk i siffror: vilka språk talas och av hur många?*, Stockholm: Morfem.

Reuter, M., 1971, 'Vokalerna I finlandssvenskan – En instrumentell analys och ett försök till systematisering enligt särdrag', *Studier i nordisk filologi* 58, 240–49.

Söderström, P., Horne, M. and Roll, M., 2016, 'Word accents and phonological neighbourhood as predictive cues in spoken language comprehension', in *Proceedings of Speech Prosody 2016*, pp. 45–48, Boston.

Söderström, P., Horne, M., Frid, J. and Roll, M., 2016, 'Pre-Activation Negativity (PrAN) in brain potentials to unfolding words', *Frontiers in Human Neuroscience* 10. https://doi.org/10.3389/fnhum.2016.00512

Söderström, P., Roll, M. and Horne, M., 2012, 'Processing morphologically conditioned word accents', *Mental Lexicon* 7, 77–89. https://doi.org/10.1075/ml.7.1.04soe

Soto-Faraco, S., Sebastian-Galles, N. and Cutler, A., 2001, 'Segmental and suprasegmental mismatch in lexical access', *Journal of Memory and Language* 45, 412–32. https://doi.org/10.1006/jmla.2000.2783

Sundström, S., Samuelsson, C. and Lyxell, B., 2014, 'Repetition of words and nonwords in typically developing children: The role of prosody', *First Language* 34(5), 428–49.

Thorén, B., 2003, 'Can V/C-ratio alone be sufficient for discrimination of V:C/VC: in Swedish? A perception test with manipulated durations', *Proceedings from Fonetik 2003*, Umeå University.

Thorén, B., 2008, *The priority of temporal aspects in L2-Swedish prosody: Studies in perception and production*, PhD thesis, Stockholm University.

Zhou, X. and Marslen-Wilson, W., 1994, 'Words, morphemes and syllables in the Chinese mental lexicon', *Language and Cognitive Processes* 9(3), 393–422. https://doi.org/10.1080/01690969408402125

Åsa Abelin is a Professor of General Linguistics at the Department of Philosophy, Linguistics and Theory of Science, University of Gothenburg. Her research interests lie within phonetics and psycholinguistics, in the areas of second-language acquisition, emotional prosody and iconicity in language. In particular, she has studied intelligibility, comprehensibility and attitudes to non-native accent as well as perception of emotional prosody, and sound symbolism and onomatopoeia in Swedish.

Bosse Thorén is a senior lecturer at Dalarna University, Sweden, and he also gives lectures to teachers of Swedish as a second language at seminars and workshops. His PhD at Stockholm University was about the priority of temporal aspects of the Swedish prosody. His research interests lie in describing Swedish and English prosodies and investigating the intelligibility of foreign-accented speech in relation to second-language learners. He seeks to suggest phonological core features of Swedish and English critical for crosscultural communication.