

Kuronen, M., P. Lintunen & T. Nieminen (toim.) 2017. Näkökulmia toisen kielen puheeseen – Insights into second language speech. AFinLA-e. Soveltavan kielitieteen tutkimuksia 2017 / n:o 10. 214–233.

**Mikko Kuronen & Elina Tergujeff**

University of Jyväskylä

## **L1 listeners' perception of L2 pronunciation: effect of prosody on accentedness ratings in Swedish**

This paper reports on a study in which L1 speakers (n = 53) of Swedish rated the accentedness of read-aloud sentences produced by L2 speakers (n = 6) of Swedish. The aim of the study was to investigate the effect of certain prosodic features on the accentedness ratings. These features include Swedish tone accents 1 and 2 and deviant speech rhythm (e.g. lack of primary stress, and realization of multiple stresses). The results reveal that utterances with deviant speech rhythm yielded perceptions of stronger foreign accent than utterances with target-like rhythm. As for the tone accents, their effect on the scores were non-existent or only marginal. As missing tone accents have been previously found not to compromise intelligibility, the results of the present study reinforce the conclusion according to which emphasis in teaching should be put on prosodic features other than tone accents, especially rhythm.

**Keywords:** pronunciation, prosody, accentedness, tone accents, rhythm

## 1 Introduction

Second language (L2) pronunciation is mainly studied by acoustic analysis and auditory tests. The acoustic characteristics of L2 Swedish have been studied concerning vowels (Hertteli 2015; Kuronen 2016), consonants (Kokkonen 2016; Kuronen 2016), tone accents (Kaiser 2011; Tronnier & Zetterholm 2014; Hed 2014; Kuronen et al. 2016), and rhythm and utterance intonation (Kuronen & Tergujeff 2018). Acoustic studies give us detailed information on what L2 pronunciation is like in comparison to first language (L1) pronunciation. In turn, auditory tests reveal how listeners experience L2 pronunciation: how comprehensible it is, how accented it is, and how pleasant it sounds. These issues are investigated by using scalar evaluations, whereas methods such as dictation are used to measure actual intelligibility (Derwing & Munro 2005). Most of the studies on L2 Swedish have been either acoustic-phonetic or they have focused on accentedness and/or comprehensibility. Thus, there is a lack of studies on relations between acoustic features and perception of accentedness.

In this study, focus is on the phonetic basis of L2 accentedness – not intelligibility or comprehensibility (cf. Derwing & Munro 1997). Our goal is to produce new knowledge on how prosodic features affect accentedness ratings by L1 listeners. We aim to analyse the perceptual importance of the main features of utterance prosody in Sweden-Swedish: tone accents, rhythm, and utterance intonation (Bruce 1977; Myrberg 2010; Riad 2014). These are difficult learning goals for L2 speakers, but there is very little knowledge on the perceptual relevance of the features. The results are applicable in teaching of pronunciation especially to advanced L2 learners of Swedish, irrespective of their L1. For example, if tone accents do not affect L1 listeners' perception of accentedness, not much effort needs to be spent by L2 learners to acquire them. Most of the previous studies on phonetic aspects of L2 perception have concerned English (see e.g. Jesney 2004; Derwing & Munro 2015). Since learning and perception of prosody is partly language-dependent, the present study on Swedish will hopefully elucidate new aspects on how L2 pronunciation is perceived.

## 2 Previous studies on perception of accentedness in L2 Swedish

There are numerous studies on how listeners perceive L2 English (e.g. Flege & Fletcher 1992; Bongaerts et al. 1997; Derwing & Munro 1997; Piske et al. 2001), and several of them are phonetically detailed (e.g. Flege 1984; Jilka 2000). The perception of L2 pronunciation in other languages has also been studied in some detail, such as in French (e.g. Flege 1987; Birdsong 2007) and

in German (e.g. Missaglia 2007). However, there are only a few studies on how Swedish listeners perceive and rate accentedness based on phonetic features. The main findings of these studies will be discussed below.

Cunningham-Andersson & Engstrand (1989) analysed the perceptual effects of different phonetic features in Finnish- and British-accented L2 Swedish, as imitated by the second author of the study. 14 phonetic features were combined to make 35 different imitations of a short text, which Swedish listeners ( $n = 35$ ) rated on a scale from 0 to 4. The focus of the analysis was on segments. Also, exaggerated quantity distinctions and mixing of tone accents were analysed. The most significant L2 features of Finnish-accented L2 Swedish were velarized /l/, unaspirated stops, trilled /r/, exaggerated quantity distinctions, and mixed tone accents (accent 2 pronounced as accent 1). Occurrence of one of these features was enough for 50% of the listeners to perceive a foreign accent, and two or more of the features increased the rating of a speech sample as foreign-accented. Identification of foreign accent as Finnish had a positive correlation with an increasing number of the above-mentioned features. Further, listeners were quite good at identifying the imitated accent as Finnish (60–100%), a finding also made by Cunningham-Andersson & Engstrand (1988a,b), Bijvoet (1996), and Boyd et al. (1999) with authentic speech material.

Kuronen & Zetterholm (2017) also used imitations. They varied concerning segmental qualities, tone accent production, utterance intonation, and rhythm. Swedish listeners ( $n = 30$ ) rated the accentedness of the imitations (a short text) on a scale from 1 to 6. Segmental deviations from L1 Swedish affected the ratings much more than deviations in utterance intonation, rhythm, and/or tone accents. Imitations with L2-like intonation and native-like segments were rated in the same way as imitations with native-like intonation and native-like segments. If segments were native-like, imitations with native-like rhythm and L2-like rhythm were rated similarly. Derwing & Munro (1997: 11) have made a similar finding concerning English: segments were the main cause for perception of foreign accent (92%), followed by grammar (46%), enunciation (mumble, 38%), prosody (23%), and speech rate (15%), as evaluated by the listeners themselves. Based on these two studies, segmental deviations seem more easily perceived as L2 features than prosodic deviations (at least in Swedish and English). This may be due to the fact that segment quality varies less audibly than prosody between different speaking styles, situations, and speakers. In other words, listeners may be more used to prosodic variation than segmental variation in their L1. Concerning Swedish, the dialectal variation might also be more noticeable in prosody than in segments (perhaps disregarding the dialects in Scania and Gotland). This too can make Swedish listeners more tolerant towards prosodic than segmental deviations when asked to rate accentedness of L2 pronunciation.

Abelin & Thorén (2015a,b) analysed the relative importance of incorrect word stress placement versus deviant tone accent. A male phonetician pronounced two-syllable words (e.g. *bilen* (Eng. *the car*), *sagan* (Eng. *the fairy tale*)) and non-words with varying stress placement and intonation. The listeners ( $n = 18$ ) were asked to decide under time pressure whether the word was a real word or a non-word in Swedish. Incorrect stress placement caused much more non-word answers than a mispronounced tone accent (accent 2 as accent 1, the opposite was not analysed). Thus, word stress placement was clearly more important for word recognition than tone accent, so it should also affect the perceived accentedness much more than tone accent.

To our knowledge, there are no other studies on the connection between phonetic features and perceived accentedness in Swedish than the ones summarised above. Taken together, the results suggest that segments are more crucial for perception of accentedness than prosody, as long as the deviations are at a phonetic level, i.e. non-phonemic. In other words, listeners seem to be more sensitive to segmental than prosodic deviations when rating accentedness. Of course, prosodic deviations at phonological level (such as incorrect word stress placement) are decisive for perception of accentedness. The speech materials in Cunningham-Andersson & Engstrand (1989) and Kuronen & Zetterholm (2017) were imitations, and since the findings concerning prosodic features in particular have not been investigated in other studies on L2 Swedish, we set our focus solely on prosody in this study. Authentic L2 speech was used as speech material, and the experimental setup was developed from Cunningham-Andersson & Engstrand (1989) and Kuronen & Zetterholm (2017).

Before we advance to the empirical part, it is appropriate to state that we think that a communicatively functional pronunciation (cf. Levis 2005) is the first priority for L2 speakers. However, the phonological aspects of L2 may not be sufficient learning goals for all advanced learners. If that was the case, the pronunciation of e.g. *kultur* (Eng. *culture*) as [kul'tu:ːr] and *kaka* (Eng. *biscuit*) as ['ka:k:a] (typical pronunciations for L1 Finnish-speaking learners of Swedish) would be adequate goals instead of [kəl'tʰɥ:ːɹ] and ['kʰɔ:k:aː] (L1 Swedish). Yet, [kul'tu:ːr] and ['ka:k:a] would certainly give an impression of a strong foreign accent. Learning goals vary depending on L2 speakers' ambitions and needs: [kul'tu:ːr] and ['ka:k:a] should be well intelligible in most contexts and can perhaps serve as basic learning goals for L2 speakers. Still, aspiration, vowel quality, post-vocalic consonant duration (in *kaka*), and/or tone accents must be adjusted/acquired if an L2 speaker aims higher than just to make him- or herself understood. This is true especially when dealing with L2 speakers who wish to acquire comfortable intelligibility or native-like pronunciation. Comfortable intelligibility is important for both professional and non-professional language users because a strong foreign accent will put more

strain on the listener, and can affect listeners' attitude toward the speaker (Mennen 2007; Torstensson 2010; Boyd & Bredänge 2013). Comfortable intelligibility can also improve an L2 speaker's self-confidence as a language user.

### 3 Empirical part

In the following sections, we will describe our aims and research questions (3.1), speech material (3.2), and methods (3.3).

#### 3.1 Aims

The present paper is based on acoustic studies conducted by the authors (Kuronen 2015; Kuronen & Tergujeff 2018). The results of these studies show that many L2 speakers of Swedish can develop prosodically during phonetic training: many participants learned to produce tone accents, their utterance intonation became more native-like, they learned to use duration as a stress correlate, and they acquired a native-like rhythm. Here, our aim is to find out how important the attainment of these features is for L1 Swedish listeners. Segmental qualities are not analysed since they have shown to be pivotal in previous studies (see Chapter 2). The following research questions are addressed:

1. Do tone accents affect foreign accentedness ratings by L1 Swedish listeners? Are tone accents 1 and 2 equally important or unimportant? Is native-like tone accent 2 more important in compounds than in simple words?
2. Does mixing up tone accents affect foreign accentedness ratings by L1 Swedish listeners?
3. Does lack of clear primary stress and tendency not to distress syllables outside the primary stress affect foreign accentedness ratings by L1 Swedish listeners?
4. Do multiple stresses affect foreign accentedness ratings by L1 Swedish listeners?

Earlier knowledge on the effect of tone accents is insufficient. According to Cunningham-Andersson & Engstrand (1989), mixing of tone accents has a negative perceptual effect, but lack of tone accents was not analysed in their study. Still, lack of tone accents is probably a more common trait in L2 Swedish than mixing of them (Kuronen 2015). In the present study, both lack of tone accents and mixing of them were analysed. Our assumption was that mixing of

tone accents has a negative perceptual effect, while lack of them has merely a marginal effect due to listeners' tolerance for dialectal and/or situational variation in prosody. However, one could expect that a native-like tone accent 2 would have some positive effect due to its distinct tonal gesture.

Tone accent 2 has a connective function in compounds (Elert 1981; see also Riad 2015). Compounds are pronounced in Standard Swedish (Central-Swedish, Riad 2014) with two stressed syllables and a high+low+high (H\*LH) contour with H in the stressed (= long) syllables, e.g. *lastbilschaufför* (Eng. *lorry driver*) ['la`s:tbi:lsʃaʊfœ:´ɹ]. Due to two stressed syllables in compounds, tone accent 2 might be more important for native-like pronunciation in compounds than in simple words (e.g. *kappa*, Eng. *coat*). If listener ratings are similar irrespective of intonation in the primary stressed word, it would support the idea of tone accents not being a purposeful learning goal for L2 speakers of Swedish.

Research questions 3 and 4 have been studied earlier only by Kuronen & Zetterholm (2017) with imitations of a text sample (see Chapter 2). The L2 speaker was highly fluent in prosody, and therefore the imitations might not be representative for a more ordinary L1 Finnish-speaking learner of Swedish. Also, only native-like rhythm and rhythm with lack of stress(es) were studied. Here, three different rhythmic patterns are studied: (i) native-like rhythm with primary stress and distressing of syllables (= a tendency to shorten them) in the unstressed words in the utterance, (ii) lack of primary stress and no distressing of lexical (= phonological) stresses, and (iii) multiple stresses. Two last-mentioned deviations are typical for L2 speakers despite their L1 (Rasier & Hiligsmann 2007; on Finnish-Speakers' Finland-Swedish see Kautonen 2017). As a consequence of the rhythmic differences, the utterance intonation varies as follows: (i) in utterances with native-like rhythm, the contour is non-descending with the highest  $f_0$  peak in the primary-stressed syllable, (ii) in utterances with lacking primary stress, the contour is descending after the initial  $f_0$  peak, and (iii) in utterances with multiple stresses, there are several distinct  $f_0$  peaks (see Figures 1–5 in section 3.3).

Knowledge on how listeners rate different rhythmic patterns is important when teaching pronunciation to L2 speakers. Native-like rhythm is considered an important learning goal for L2 speakers of Swedish (e.g. Bannert 2004), but knowledge on perception of different deviating patterns is scarce.

### 3.2 Speech material

The speech material consisted of read-aloud sentences; mainly statements ( $n = 13$ ), but a few questions ( $n = 3$ ) were also included. The utterances had (or should have had) primary stress towards the end, and with tone accent 1 or 2 depending on the utterance. The utterances were 1.5 to 2.5 seconds long,

and they were prosodically quite simple, meaning they did not vary much regarding primary stress placement, tone accents, rhythm, and utterance intonation, as pronounced by native-speaker controls. Below, five example utterances from the speech material are given. The usual primary stress placement in L1 pronunciation is marked with blue.

- |  |  |
|--|--|
| 1. Ska vi åka <i>ut</i> ikväll?              | Are we going <i>out</i> tonight?             |
| 2. Jag ska köpa en <i>matta</i> till pappa.  | I will buy a <i>carpet</i> to my father.     |
| 3. Vi ska till <i>stugan</i> ikväll.         | We are going to the <i>cottage</i> tonight.  |
| 4. Jag vill köpa en <i>kappa</i> till mamma. | I want to buy a <i>coat</i> for (my) mother. |
| 5. Det kan <i>blåsa</i> ikväll.              | It might get <i>windy</i> tonight.           |

We preferred to use authentic L2 speech instead of synthesized speech for two reasons. Firstly, we had collected a large L2 speech corpus, which enabled us to find utterances with appropriate phonetic characteristics; tokens with prosody of the wanted types. Secondly, synthetic editing of rhythm and utterance intonation can easily make utterances sound unnatural. Use of read-aloud sentences instead of (longer) samples of free speech is based on the assumption that in read-aloud sentences listeners can more easily focus on the accentedness, because the contents have been controlled (e.g. no grammatical errors). Influence of the speaker's voice can be assumed weaker in short utterances than in longer speech samples. In longer samples, accentedness may be easier to rate, but phonetic factors behind the ratings are very difficult to distinguish.

The L2 speakers ( $n = 6$ , all female) who produced the test utterances were high-intermediate/advanced speakers of Swedish with Finnish as their L1. L1 speakers ( $n = 2$ , females from Stockholm) were used for reference. The L2 speakers were 21–36 years old, while the L1 speakers were 24 and 25 years old. The L2 speakers had studied Swedish for at least six years as part of their secondary education before their current studies of Swedish as a major or minor subject at a Finnish university. They took part in a course in Swedish pronunciation and oral skills, and their pronunciation was recorded in a language lab before and after the course. The recordings were saved in a sound file format WAV. The recordings of the L1 speakers were made in a recording studio and saved in a sound file format AIFF. Both file formats used in the collection of speech materials are uncompressed formats. The speech material for the present study was chosen from these recordings. The quality of the recordings for the L1 and L2 speakers was similar.

### 3.3 Auditory test

We chose 30 utterances for the auditory test from a larger material of ca. 250 utterances. 24 of the utterances were pronounced by L2 speakers, 3 utterances by L1 speakers, and 3 utterances were presented twice to the listeners

to test the reliability of the ratings. We used the same L2 speakers in as many of the utterances as possible: 17 of 24 L2 utterances were pronounced by only three speakers. All utterances in which the effect of tone accents was analysed were pronounced by these three speakers. Most of the utterances occurred twice in the auditory test representing different prosodic categories. Overall, the selection of L2 utterances was based on three criteria: (i) the segments are as native-like as possible so that defects in them would affect the ratings as little as possible, (ii) there are no pauses, hesitation, and/or corrections in the utterances, and (iii) in the utterances with tone accent, there is a clear primary stress (prominence level 2, Myrberg 2010), meaning that this stressed word should have either tone accent 1 (L\*H) or tone accent 2 (H\*LH), as was the case in the native speakers' pronunciation.

The L2 speakers have some segmental deviations from L1 pronunciation. These deviations may affect the ratings, but they are similar in all analysed prosodic categories, because (i) the speakers are the same in the vast majority of the utterances, (ii) the speakers have the same kind of segmental deviations, primarily caused by their L1 Finnish, and (iii) we have chosen the utterances so that segmental differences between them are as minor as possible. Thus, the outcome of the auditory test should substantially depend on the prosodic character of the utterances, not the segmental quality in them.

The utterances were chosen as follows: The first author listened to the larger material, and chose 40 suitable utterances, which were presented to the second author. Only when both authors categorised a certain utterance belonging to one of the eight categories below, the utterance was used in the test. In other words, the material selection was a result of careful negotiation between the authors. The prosodic categories that the chosen utterances demonstrate are:

1. correct primary stress and rhythm, no tone accent 1 (H\*L)
2. correct primary stress and rhythm, tone accent 1 (L\*H)
3. correct primary stress and rhythm, no tone accent 2 (H\*L)
4. correct primary stress and rhythm, tone accent 2 (H\*LH)
5. correct primary stress and rhythm, mix-up of tone accents (accent 1 pronounced as accent 2, and vice versa)
6. correct primary stress and rhythm, no tone accent 2 in a primary-stressed compound with four or more syllables
7. L2-like rhythm with lack of distinct primary stress and no distressing of the phonologically stressed syllables in the utterance

8. L2-like rhythm with multiple stresses, short stress groups, no distinct primary stress
9. L1 Swedish speakers' pronunciation.

We used three utterances per category in the auditory test (3 x 8 L2 prosodic categories + 3 L1 utterances + 3 control utterances by L2 speakers = total of 30 utterances). This was done so that minor differences in segment quality and other possible individual speaker qualities would not skew the results. No more than 30 utterances were used so that the auditory test would not take more than 15 minutes to finish.

In Figures 1–5, segment quality and reduction (cf. transcription), syllable durations (stress and rhythm), and intonation in five utterances belonging to five different prosodic categories are presented as examples. In these figures all the varied prosodic features are illustrated: L1- and L2-like utterance intonations (Figures 1 and 4, respectively), L1- and L2-like tone accent productions (Figures 2 and 3, respectively), and L1- and L2-like rhythmic patterns (Figures 1 and 5, respectively). The analyses have been conducted with Praat (Boersma & Weenink 2016).

The listeners were asked to rate the accentedness of each utterance on the following scale:

- 6 = native-like
- 5 = really good pronunciation, almost native-like
- 4 = quite good pronunciation with an audible foreign accent
- 3 = good pronunciation with a clear foreign accent
- 2 = quite poor pronunciation with a strong foreign accent
- 1 = very poor pronunciation with a very strong foreign accent

In most of the earlier studies, 5 to 9 point scales have been used (Jesney 2004). Here, we used a 6 point scale because it worked well in a similar study made earlier by the first author (Kuronen & Zetterholm 2017). Further, we wanted to obtain our results in a form that is easy to compare with the previous study. A scale with 7 to 10 points would probably bring forth even more subtle differences, but the description of the scale is more difficult with a larger scale, and we wanted to describe the scale points for precision. Because the L2 speakers had quite good pronunciation skills, we anticipated that the listeners would mainly use the upper part of the scale.

The auditory test was done with an online survey tool SoGoSurvey. Before the test, the listeners were told that (i) the context of the test was learning

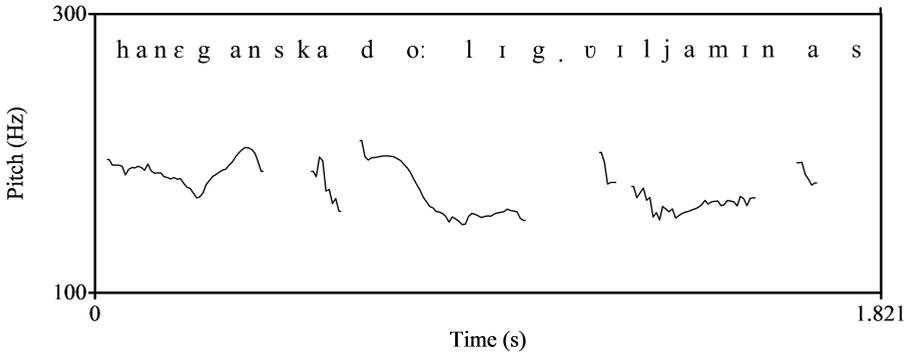


FIGURE 1. L2 speaker's pronunciation of *Han är ganska dålig, vill jag minnas* (Eng. *He is quite bad, I recall*). The utterance belongs to group 3: correct primary stress and rhythm, no tone accent 2 in primary stress (H\*LL instead of H\*LH). The most obvious prosodic L2 trait is lack of tone accent 2 despite a clear primary stress in *dålig* (Eng. *bad*).

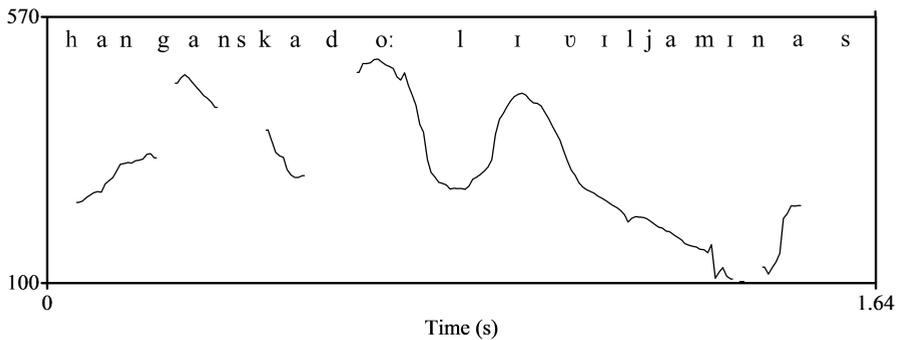


FIGURE 2. L1 speaker's pronunciation of *Han är ganska dålig, vill jag minnas* (Eng. *He is quite bad, I recall*). The most obvious prosodic difference compared with Figure 1 is tone accent 2 (H\*LH) in the primary stressed *dålig* (Eng. *bad*).

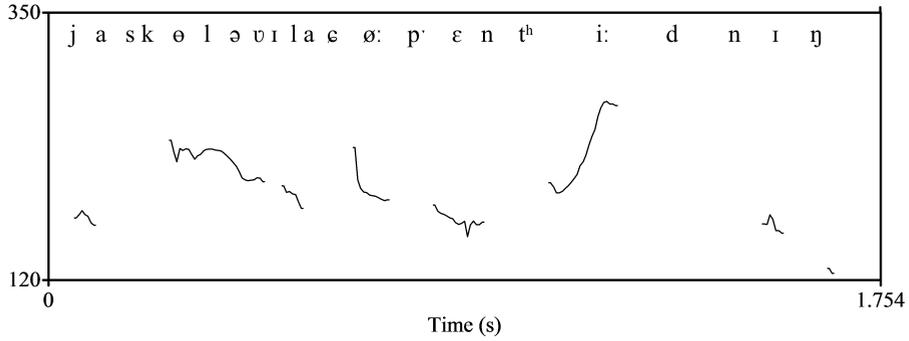


FIGURE 3. L2 speaker's pronunciation of *Jag skulle vilja köpa en tidning* (Eng. *I would like to buy a newspaper*). The utterance belongs to group 5: correct primary stress and rhythm, mixed tone accents (tone accent 2 realized as tone accent 1) in the primary stressed *tidning* (Eng. *newspaper*).

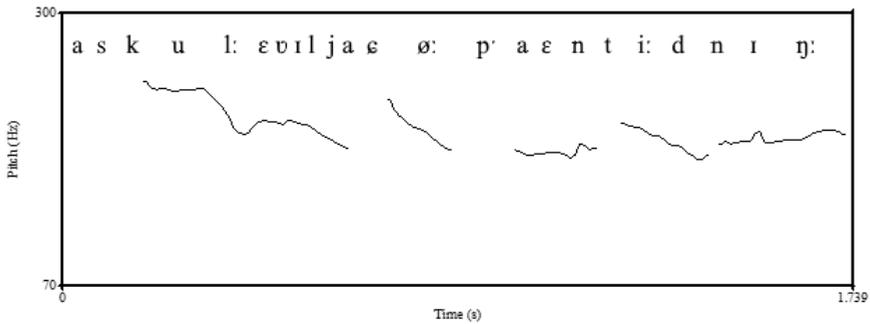


FIGURE 4. L2 speaker's pronunciation of *Jag skulle vilja köpa en tidning* (Eng. *I would like to buy a newspaper*). The utterance belongs to group 7: L2 like rhythm with lack of a distinct primary stress and no distressing of the phonologically stressed syllables in the utterance.

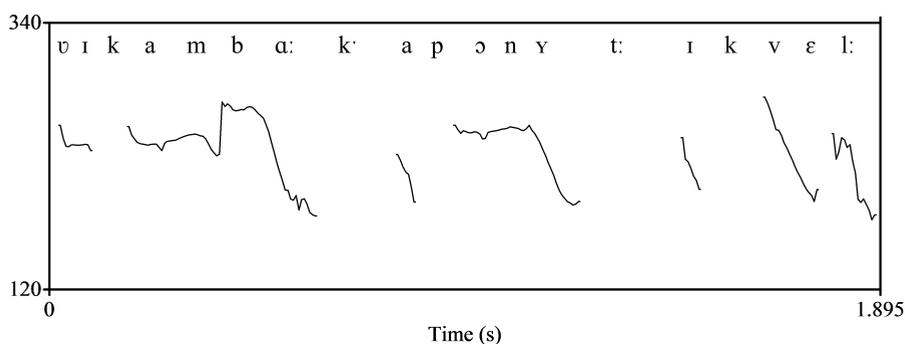


FIGURE 5. L2 speaker's pronunciation of *Vi kan baka på nytt i kväll* (Eng. *We can bake again tonight*). The utterance belongs to group 8: L2 like rhythm with multiple stresses, short stress groups, no distinct primary stress. Here, the stressed syllables are *baka* (Eng. *bake*), *pånytt* (Eng. *again*), and *ikväll* (Eng. *tonight*).

of Sweden-Swedish pronunciation, (ii) they will hear 30 utterances, (iii) their only task is to rate the accentedness of each utterance on a given scale, and (iv) the utterances have no major defects in pronunciation, e.g. completely incorrect speech sounds. The last-mentioned information was given to offer the listeners prior knowledge about the speakers' advanced proficiency in Swedish. The rating scale and the description of it were shown on the screen below the sound icon for each utterance. The utterances were presented to the listeners in a randomised order. The listeners were instructed to listen to each utterance as many times as they wanted. Statistical analyses were done using a reliability alpha (Cronbach's  $\alpha$ ) and independent samples t-tests. Effect sizes of the statistical significances were calculated using Cohen's  $d$ .

The listeners ( $n = 53$ ) were adults with Sweden-Swedish as their L1 and with no reported hearing problems nor studies in the Finnish language. They were invited through the authors' contacts in Sweden to take the listening test using their personal computers and headphones. No previous experience in accentedness ratings nor phonetic expertise was required. The majority of the listeners had an East Central Swedish pronunciation (cf. Standard Swedish; Riad 2014). They were not asked to rate the utterances in regards to any specific variety of Sweden-Swedish, because such information probably would have confused them: if e.g. East Central Swedish or Standard Swedish had been mentioned in the instructions, some of the listeners might have thought that they are unable to define these (regional) varieties and/or unable to do the ratings for some other reason. We knew before the test that most of the listeners resided in areas where East Central Swedish is spoken, and we thought that it was natural to them to rate the utterances with regard to East Central Swedish, not e.g. the variety spoken in Scania. Further, L2 speakers had Stan-

ard Swedish as their target variety, which could be heard in segment qualities – a circumstance that also directed the listeners to rate the utterances with regard to Standard Swedish.

## 4 Results and analyses

In the following sections, we will report the results. Scale and listener reliability are commented in section 4.1. The ratings by prosodic feature and by speech sample are reported and analysed in section 4.2.

### 4.1 Scale and listener reliability

When rating the 24 test samples pronounced by L2 speakers, mean scores given by the raters ( $n = 53$ ) varied from 2.54 to 5.79. Thus, some raters were much stricter in their scores than others. However, the listeners were consistent in their ratings: a reliability alpha (Cronbach's  $\alpha$ ) for all ratings, including native-speaker controls and check samples, was excellent ( $\alpha = 0.949$ ), indicating very high rater consistency and reliability of the scale.

Native-speaker controls and check samples served well in their purpose. Raters made a clear distinction between L2 speakers and native speakers by scoring the native-speaker samples substantially higher than the L2 speaker samples. Mean score for L2 speakers was 4.18 (SD 1.10), whereas it was 5.98 (SD 0.14) for native speakers. Independent samples t-test yielded a statistically significant difference between scores received by L2 speakers and native speakers ( $p = 0.000$ ), and the effect size was very large ( $d = 2.91$ ). On the contrary, no statistically significant differences were found between the two rounds of ratings for the three check samples ( $p = 0.44$ ,  $p = 0.61$ ,  $p = 0.55$ ).

### 4.2 Accentedness ratings

The mean scores for each prosodic feature category are presented in Figure 6. The results reveal that mean scores for the six different tone accent categories were between 4.20 and 4.73, whereas deviant rhythm (lack of primary stress or multiple stresses) received mean scores closer to 3 (3.48 and 3.18, respectively). No major differences were found in mean scores for tone accent categories, but the mean score for target-like accent 2 was the highest (4.73). Mixing up accent 1 and accent 2 did not affect the ratings. The accent mix-up category received a mean score of 4.57, which is almost as high as the mean score for target-like accent 2 and higher than the mean score for target-like accent 1 (4.20). Standard deviation for target-like accent 2 was 0.89, whereas it was between 0.95 and 1.01 for the other categories.

When scores for samples presenting no tone accents and target-like accents were subjected to an independent samples t-test, no statistically significant difference was found between no accent 1 and target-like accent 1 ( $p = 0.53$ ). This was not the case with accent 2, for which the t-test resulted in a statistically significant difference ( $p = 0.009$ ) between missing tone accent and target-like productions. However, the effect size was small ( $d = 0.29$ ). When scores for target-like accent 2 were compared to scores of the accent mix-up category, no statistically significant difference was found ( $p = 0.14$ ). Similarly, the difference between target-like accent 2 and no accent 2 in compounds was not statistically significant at the 0.05 level ( $p = 0.06$ ).

The outcome that native-like tone accents have no or only a very marginal perceptual effect is in line with the findings in Kuronen & Zetterholm (2017). Still, it is surprising that not even missing accent 2 in compounds got lower ratings than correct tone accent 2. Further, it is curious that mixing up tone accents did not affect the ratings.

The results concerning tones might be caused by two factors: (i) listeners are very tolerant for variation in word intonation in both L1 and L2 Swedish, and/or (ii) they hear that the speakers are non-native in Swedish, and therefore ignore the tonal gesture in the primary-stressed word. The latter would mean that once an L1 listener hears that a speaker is an L2 speaker, the tonal gesture in the stressed word does not affect the rating. Further, the L2 speakers in this study may have slight differences in their productions of tone accents if compared to native speakers. Yet, these speakers had somewhat native-like pronunciation, and if tone accents have no effect as produced by them, this should apply to many other L2 speakers as well.

The results concerning tone accents could also depend on the fact that listeners perceived the stressed words in L2 speakers' utterances as pronounced as prominence level 1 instead of prominence level 2. As prominence level 1, tone accents are pronounced as HL\* (tone accent 1) and H\*L(L) (tone accent 2) – in other words, with a descending contour like they were produced by L2 speakers in some of the utterances. Even with this aspect taken into account, it remains to be answered why mixing of tone accents did not have a negative effect on the ratings.

All six prosodic categories with native-like rhythm received much higher ratings than the two categories with rhythmic deviations (Figure 6). The difference between native-like rhythm categories and the two deviant rhythm categories was statistically significant ( $p = 0.000$ ), and the effect size was large ( $d = 1.13$ ). It is not surprising that multiple stresses had the strongest negative effect on the ratings, because this pattern differs from native-like rhythm by adding a disturbing component and making stress groups short. Somewhat higher but still relatively low scores were given to lacking primary stress and no distressing of phonologically stressed syllables (which is a common pro-

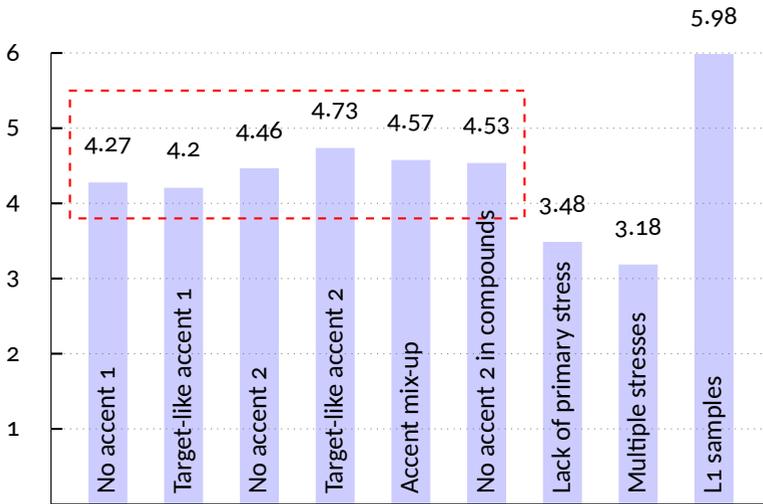


FIGURE 6. Mean scores by prosodic feature, including scores received by L1 samples. Utterances with native-like rhythm are marked with a rectangle.

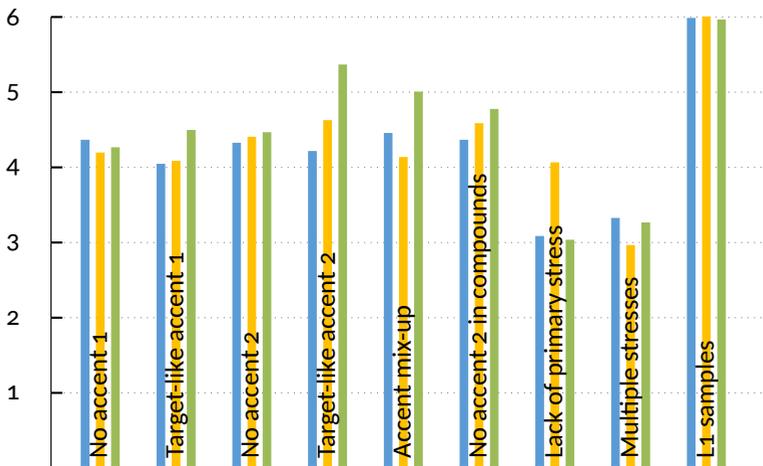


FIGURE 7. Mean scores by speech sample, including scores received by L1 samples.

cess in Swedish; Figure 6). The difference between multiple stresses and lack of primary stress is statistically significant ( $p = 0.009$ ). However, the effect size was small ( $d = 0.29$ ). In Kuronen & Zetterholm (2017), lack of primary stress had a very marginal negative effect on accentedness ratings. The difference between the findings here and in Kuronen & Zetterholm is most probably due to the fact that the speaker in Kuronen & Zetterholm was more native-like in other aspects of pronunciation than the L2 speakers in utterances with lacking primary stress in this study; his articulation rate and fluency were better. In that kind of pronunciation, lack of primary stress can be considered merely as a situational variation by L1 listeners, not as a trait of L2 pronunciation.

Ratings by feature were formed of scores received by three speech samples per feature, i.e. 159 scores in total. Figure 7 presents mean scores for each speech sample. The figure shows that not all speech samples within one category scored equally. Most variation was found within the following categories: target-like accent 2, no accent 2 in compounds, and lack of primary stress. This is probably due to the minor segmental differences between the utterances. In addition, a somewhat lower speech rate and  $f_0$  may have lowered some scores. Aside from three utterances, the coherence of the ratings is quite evident. Also, the native speakers were recognized as native speakers with a very high precision, presumably based on both prosody and segments in their utterances.

## 5 Summary and concluding remarks

The purpose of this study was to find out how important tone accents and different rhythmic patterns are for Swedish listeners' ( $n = 53$ ) ratings of L2 accentedness. The speech material consisted of 1.5 to 2.5 sec. long utterances with different prosodic characteristics. L1 speakers were used for reference. The results are summarized as follows:

- (i) Tone accents play no or only a very marginal role for L1 listeners' ratings.
- (ii) Not even mixing of tone accents had a negative perceptual effect.
- (iii) Utterances with compounds and without correct tone accent got ratings as high as utterances with correct tone accent 2.
- (iv) Rhythm with a clear primary stress and distressing of syllables outside primary stress got high ratings.
- (v) Rhythm with multiple stresses got lowest ratings of all.
- (vi) Lack of primary stress rendered also relatively low ratings.

Kuronen & Zetterholm (2017) found out that lack of tone accents did not have a negative perceptual effect even in longer speech samples (21–24 sec.). The speaker was a near native-like speaker of Swedish. The results here and in Kuronen & Zetterholm suggest that the acquisition of tone accents does not seem to make L2 pronunciation more native-like. Thus, even advanced L2 speakers of Swedish should focus in their learning on aspects of pronunciation other than tone accents.

In contrast to our results, Cunningham-Andersson & Engstrand (1989) found out that the mixing of tone accents had a clear negative perceptual effect. They used imitations by a native speaker. Further, the material consisted of a short text with multiple occurrences of mixed tones. Both of these factors can cause the difference between the findings here and in Cunningham-Andersson & Engstrand's study.

Concerning rhythm, our findings suggest that a clear primary stress and distressing of syllables outside the primary stress is an important feature for Swedish listeners. In Kuronen & Zetterholm (2017), lack of primary stress had a very marginal negative effect on accentedness ratings. The difference between the findings here and in Kuronen & Zetterholm is most probably due to the difference in articulation rate – a factor found to affect accentedness (Derwing & Munro 2015). Thus, the present study implies that articulation rate and mean *f<sub>0</sub>* may also play a role when L1 Swedish listeners rate accentedness in L2 Swedish. These factors should be addressed in future studies.

Segmental deviations were not a variable in our experimental setup. Yet, the results show indirectly that L1 listeners are very sensitive to them: the highest score given to an L2 utterance was 5.36; all other L2 utterances were rated below 5.0 despite the native-like prosody in some of them. Imitations with some deviating prosodic characteristics but native-like segments got considerably higher ratings (5.60–5.63) in Kuronen & Zetterholm (2017). These findings suggest that segments are very important if the goal of an L2 speaker is to acquire as unaccented pronunciation as possible.

The results concerning tone accents should not be interpreted as if they played *no role* in native-like pronunciation. Namely, there seems to be a connection between the acquisition of tone accents and other aspects of intonation such as overall contour, *f<sub>0</sub>* range, and mean *f<sub>0</sub>* (Kuronen & Tergujeff 2018). Thus, the learning of tones can facilitate the learning of other aspects of intonation, but the perceptual effect with native listeners still seems to be non-existing or at least very marginal.

## Acknowledgements

This study is part of a research project called *Fokus på uttalsinläringen med svenska som mål- och källspråk* (<https://www.jyu.fi/fokus>). The project is funded by The Society of Swedish Literature in Finland (SLS), which also supported Tergujeff's work with an additional research grant.

## References

- Abelin, Å. & B. Thorén 2015a. The relative perceptual weight of two Swedish prosodic contrasts. In E. Babatsouli & D. Ingram (eds) *Proceedings of the international symposium on monolingual and bilingual speech 2015*. Chania: Institute of Monolingual and Bilingual Speech, 1–7.
- 2015b. What affects recognition most: wrong word stress or wrong word accent? In M. Svensson Lundmark, G. Ambrazaitis & J. van de Weijer (eds) *Proceedings from Fonetik 2015: Lund, June 8–10 2015*. Working Papers 55. Lund: University of Lund, 7–10.
- Bannert, R. 2004. *På väg mot svenskt uttal*. Lund: Studentlitteratur.
- Bijvoet, E. 1996. Attityder till några varieteter av talad svenska. In M. Thelander, L. Elmevik, B.-L. Gunnarsson & B. Melander (eds) *Samspel och variation: språkliga studier tillägnade Bengt Nordberg på 60-årsdagen*. Uppsala: Department of Scandinavian Languages, University of Uppsala, 55–66.
- Birdsong, D. 2007. Nativelike pronunciation among late learners of French as second language. In O.-S. Bohn & M. J. Munro (eds) *Language experience in second language speech learning*. Amsterdam: John Benjamins, 99–116. DOI: 10.1075/lllt.17.12bir.
- Boersma, P. & D. Weenink 2016. PRAAT. <http://www.fon.hum.uva.nl/praat/>.
- Bongaerts, T., C. van Summeren, B. Planken & E. Schils 1997. Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19 (4), 447–465. DOI: 10.1017/S0272263197004026.
- Boyd, S., Å. Abelin & B. Dorriots 1999. Attitudes to foreign accents. In R. Andersson, Å. Abelin, J. Allwood & P. Lindblad (eds) *Fonetik 99: proceedings from the twelfth Swedish phonetics conference*. Gothenburg: University of Gothenburg, 31–35.
- Boyd, S. & G. Bredänge 2013. Attityder till brytning: exemplet utländska lärare i svenska skolor. In K. Hyltenstam & I. Lindberg (eds) *Svenska som andraspråk: i forskning, undervisning och samhälle*. 2nd ed. Lund: Studentlitteratur, 437–458.
- Bruce, G. 1977. *Swedish word accents in sentence perspective*. Lund: LiberLäromedel.
- Cunningham-Andersson, U. & O. Engstrand 1988a. Attitudes to immigrant Swedish: a literature review and preparatory experiments. In O. Engstrand, M. Dufberg & J. Stark (eds) *Perilus XIII*. Stockholm: Institute of Linguistics, University of Stockholm, 103–152.
- 1988b. Which foreign accents can Swedes identify? In *Papers from the second Swedish phonetics conference*. Lund: Department of Linguistics, University of Lund, 65–86.

- Cunningham-Andersson, U. & O. Engstrand 1989. Perceived strength and identity of foreign accent in Swedish. In O. Engstrand, M. Dufberg & C. Kylander (eds) *Perilus X: experiments in speech processes*. Stockholm: Institute of Linguistics, University of Stockholm, 65–86. <https://www.ling.su.se/perilus/perilus-x>.
- Derwing, T. M. & M. J. Munro 1997. Accent, intelligibility and comprehensibility: evidence from four L1s. *Studies in Second Language Acquisition*, 20 (1), 1–16. <https://www.cambridge.org/core/journals/studies-in-second-language-acquisition/article/div-classtileaccent-intelligibility-and-comprehensibilitydiv/729C15F62F9EC9A51A33EAB5C2D05ED0>.
- 2005. Second language accent and pronunciation teaching: a research-based approach. *TESOL Quarterly*, 39 (3), 379–397. DOI: 10.2307/3588486.
- 2015. *Pronunciation fundamentals: evidence-based perspectives for L2 teaching and research*. Amsterdam: John Benjamins.
- Elert, C.-C. 1981. *Ljud och ord i svenskan 2*. Uppsala: Almqvist & Wiksell.
- Flege, J. E. 1984. The detection of French accent by American listeners. *Journal of the Acoustical Society of America*, 76 (3), 692–707. DOI: 10.1121/1.391256.
- 1987. The production of “new” and “similar” phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, 47–65.
- Flege, J. E. & K. L. Fletcher 1992. Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America*, 91 (1), 370–389.
- Hed, A. 2014. Perception and production of Swedish word accents by Somali L1 speakers. In M. Heldner (ed.) *Proceedings from FONETIK 2014: Stockholm, Sweden, 9–11 June 2014*. Stockholm: Department of Linguistics, University of Stockholm, 105–110.
- Hertteli, H. 2015. *Inlärnin g av sverigesvenska vokaler hos universitetsstudenter med finska som S1*. MA thesis. University of Jyväskylä. <http://urn.fi/URN:NBN:fi:jyu-201511213759>.
- Jesney, K. 2004. *The use of global foreign accent rating in studies of L2 acquisition*. Calgary: Language Research Centre, University of Calgary. <http://arts.ucalgary.ca/lrc/research/reports> [accessed 5/30/2017].
- Jilka, M. 2000. Testing the contribution of prosody to the perception of foreign accent. *New Sounds*, 4, 199–207.
- Kaiser, R. 2011. Do germans produce and perceive the swedish word accent contrast? a cross-language analysis. *Tal, musik, hörsel – Quarterly Progress and Status Report*, 51 (1), 93–96.
- Kautonen, M. 2017. Finskspråkiga talares intonation av finlandssvenska i påståendeyttranden i fritt tal. *Folkmålsstudier*, 55, 31–60. <http://urn.fi/URN:NBN:fi:jyu-201710314099>.
- Kokkonen, A.-M. 2016. *Avancerade finskspråkiga inlärares uttal av svenska och tyska frikativor*. MA thesis. Jyväskylä: University of Jyväskylä. <http://urn.fi/URN:NBN:fi:jyu-201612145097>.
- Kuronen, M. 2015. Tonaccenterna i avancerade finska inlärares svenska: en akustisk studie. *Nordand – Nordisk tidsskrift for andrespråksforskning*, 10 (1), 53–80. <http://urn.fi/URN:NBN:fi:jyu-201508042611>.
- 2016. Avancerade finskspråkiga inlärares uttal av segment i sverigesvenska. *Puhe ja kieli*, 36 (3), 175–196. <https://journal.fi/pk/article/view/59009/20417>.

- Kuronen, M. & E. Tergujeff 2018. Second language prosody and its development: connection between different aspects. *The Language Learning Journal*. (Online ahead of print). <https://www.tandfonline.com/doi/full/10.1080/09571736.2018.1434228>.
- Kuronen, M., R. Ullakonoja & M. Kautonen 2016. Inläringen av de svenska tonaccenterna hos finska S2-talare: automatiseras uttalet? *Språk och stil*, 26, 161–194. <http://urn.fi/URN:NBN:fi:jyu-201801221305>.
- Kuronen, M. & E. Zetterholm 2017. Olika fonetiska drags relativa betydelse för upplevd inföddlighet i svenska. *Nordand – Nordisk tidsskrift för andrespråksforskning*, 12 (2), 134–156. <http://urn.fi/URN:NBN:fi:jyu-201712194777>.
- Levis, J. 2005. Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39 (3), 369–377. DOI: 10.2307/3588485.
- Mennen, I. 2007. Phonological and phonetic influences in non-native intonation. In J. Trouvain & U. Gut (eds) *Non-native prosody: phonetic description and teaching practice*. Berlin: Mouton de Gruyter, 53–76.
- Missaglia, F. 2007. Prosodic training of Italian learners of German: the contrastive prosody method. In J. Trouvain & U. Gut (eds) *Non-native prosody: phonetic description and teaching practice*. Berlin: Mouton de Gruyter, 237–258.
- Myrberg, S. 2010. *The intonational phonology of Stockholm Swedish*. Stockholm: Department of Scandinavian Languages, University of Stockholm. <http://urn.kb.se/resolve?urn=urn%3Anbn%3Ase%3Asu%3Adiva-38697>.
- Piske, T., I. R. A. MacKay & J. E. Flege 2001. Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, 29 (2), 191–215. DOI: 10.1006/jpho.2001.0134.
- Rasier, L. & P. Hilgsmann 2007. Prosodic transfer from L1 to L2: theoretical and methodological issues. *Nouveaux cahiers de linguistique française*, 28, 41–66.
- Riad, T. 2014. *The phonology of Swedish*. Oxford: Oxford University Press.
- 2015. *Prosodin i svenskans morfologi*. Stockholm: Morfem.
- Torstensson, N. 2010. *Judging the immigrant: accents and attitudes*. Umeå: Department of Language Studies, University of Umeå.
- Tronnier, M. & E. Zetterholm 2014. Appropriate tone accent production in L2-swedish by L1-speakers of Somali? *Concordia Working Papers in Applied Linguistics*, 5, 722–736.