Testing the Language Melody Game

An ERP, perception and production study of L2 acquisition of the Swedish word accent – grammar association.

Anna Hed

Supervisor: Mikael Roll, Andrea Schremm (co-supervisor)
Centre for Languages and Literature, Lund University

MA in Language and Linguistics, specialization in Phonetics
SPVR01 Language and Linguistics: Degree Project – Master's (Two Years) Thesis, 30 credits

June 2016

Abstract

Previous research has shown that L1 speakers use Swedish word accents to predict what linguistic material is to follow. This study investigated how L2 learners of Swedish process the word accent – suffix association. Low to intermediate level L2 learners trained this association for two weeks, using the *Language Melody Game*. Before and after training, participants took part in a pre- and post-test, in which data on production (read speech), perception and ERPs were collected. The perception test consisted of auditory stimuli with valid and invalid accent – suffix combinations, and the participants were asked about the meaning of the stimuli. After training, a pre-activation negativity (PrAN) effect increased, indicating that the participants used the accents as predictors more extensively after training. A LAN was also found for invalid stimuli, indicating that they had acquired the tone-suffix association rule. However, no P600 was found, suggesting that the participants did not reprocess the ungrammaticality, and relied more on the suffix itself. This was also seen in response times in the post-test, where invalidity did not yield longer response times, as it does for L1 speakers. The perception training also had an effect on production of word accents.

Key words: second language acquisition, ERP, word accents, prosody, PrAN, LAN, language processing

Acknowledgements

This thesis would not have been written without the help of a number of people. First of all I want to thank my supervisor Mikael Roll, and co-supervisor Andrea Schremm for guidance through the process. The rest of the neurolinguistics group at the Centre for Language and Linguistics in Lund, who helped with ideas, support and assistance in the experiment sessions, should also be mentioned here. Thanks Merle Horne, Frida Blomberg, Pelle Söderström, Mikael Novén and Otto Ewald. Thanks to Kelly Werner who recorded all the stimuli for the Language Melody Game, and all the participants who played it and willingly participated in the experiment. Thanks also to friends and family who cheered me on all the way. The work was supported by Marcus and Amalia Wallenberg Foundation (grant number 2014.0039).

Table of content

1 Introduction

Swedish word stems carry one of two tones, called word accents, associated with suffixes and morphological rules. Since the accent on the stem is associated with its following suffix, speakers with Swedish as first language (L1) are able to predict which suffix is to come by just listening to the accent on the stem. This has been shown both in behavioral studies (Söderström et al., 2012) and Electroencephalography (EEG) studies, where the event-related potential (ERP) component pre-activation negativity (PrAN) indicates that word accents pre-activate suffixes (Roll et al., 2015; Söderström et al., submitted).

However, little is known about second language (L2) processing of word accents. It has been shown, in production and perception studies, that speakers with non-tonal L1s and no training do not acquire them in Swedish (Kaiser, 2011; Tronnier & Zetterholm, 2013), but speakers with Somali L1, which is a word accent language, might do (Tronnier & Zetterholm, 2013; 2014; Hed, 2014). A study on the effect of L1 and training on Norwegian word accents showed that speakers of a tonal L1 had an advantage compared to speakers of a non-tonal L1 in a perception test (Van Dommelen & Husby, 2009). Schremm et al. (2016), however, showed in a perception study that advanced L2 learners did use word accents predictively, similarly to Swedish L1 speakers, but Gosselke Berthelsen et al. (in preparation) found that German L1 speakers with a low to intermediate level of Swedish showed no signs of predictive use of word accents when listening to them in an ERP study, as indicated by the lack of a PrAN effect. It is from here this thesis takes its standpoint. Since the early L2 participants did not show predictive use of word accents, but the more proficient ones did, the question is how this is acquired. With both a neurolinguistic and behavioral approach, this study aims to answer this question. More specifically, it aims to investigate how low to intermediate L2 learners of Swedish are affected by a two-week training period of the word accent - suffix association. And since the more proficient learners showed L1-like behavior, another question is whether their word accent processing will become more L1-like after this training.

Using similar methods as Schremm et al. (2016) and Gosselke Berthelsen et al. (in preparation) and previous studies on Swedish L1 speakers (e.g. Roll et al., 2010; 2013; 2015; Söderström et al., in press), this study tested the hypotheses that training would lead to predictive use of the word accents, and that production, perception and neural response would be expected to become more L1-like after training.

Concerning the training, a game application, currently in development, was used. This game application, the *Language Melody Game* (LMG) consists of perception training where the player trains the association between accents and suffixes. The participants played the game at home for two weeks. Before and after training, they participated in a pre- and post-test in the laboratory to collect behavioral and EEG-data. The game consists of perception training only, but this study also investigates whether playing the game also enhances production. Previous research has found that perception training is effective for L2 tone production (Wang et al., 2003a). Therefore, it is predicted that perception training will indeed enhance production.

By researching the L2 learning of word accents in Swedish, one purpose of this study is to contribute to the field of L2 learning of tonal features in a greater linguistic context. The neurolinguistic point of view is crucial to understand the nature of online L2 language processing. Furthermore, there is evidence that signs of learning can be seen earlier in neurolinguistic data than in behavioral measures (van Hell & Tokowicz, 2010). Hence, this study aims to give a good insight into the learning process, registering not only production data and behavioral response to perception, but also ERPs. Another purpose is to contribute to the area of practical applications of second language teaching of Swedish. If the game used in this study shows to be effective, it can become a useful tool for second language learners and a good aid in second language teaching. Furthermore, this study aims to lay a piece in the puzzle, which is the general understanding of L2 acquisition, and on a larger plane: the processes of the human brain.

Below follows a Background chapter (2), introducing the theories and previous research which make up the foundation of this thesis. After that, the different methodologies used in the study are presented in the Methods chapter (3). This is followed by the results from the study in the Results chapter (4), and thereafter, these are discussed in relation to previous research in the Discussion chapter (5). Lastly, some conclusions are presented, along with some directions for further research in the Conclusions chapter (6).

2 Background

2.1 Swedish word accents

As mentioned in the Introduction, Swedish is a word accent language with two word accents (Bruce, 1977). This means that every word is assigned one of two different phonological tones. The phonetic patterns of the accents differ in the regional varieties of Swedish (see Figures 1 to 4), and some varieties, such as varieties of Swedish spoken in Finland, do not make the word accent distinction at all (Bruce & Gårding, 1978). For this study, however, the Central Swedish variety will be considered, and the South Swedish variety will be briefly touched upon.

Myrberg & Riad (2015) present a prosodic hierarchy of Central Swedish, which can be used to understand the function of word accents. The hierarchy consists of different phonological levels: two levels of the prosodic word (minimal and maximal), the phonological phrase and the intonational phrase.

The minimal prosodic word ($_{\omega min}$) is the unit where stress is assigned, and only one and no more than one stress is assigned to each minimal prosodic word. Most simplex words contain one minimal prosodic word, and thereby one stress only. Compounds and some derivations consist of two minimal prosodic words, and therefore have both a primary and a secondary stress. Another characteristic of the minimal prosodic word, and a way of detecting it, is that syllabification is made within this unit, which is shown by some consonants, e.g. /h/ that will be produced as [h] in the beginning of a syllable if it is initial in the minimal prosodic word, but it gets reduced to ø or [ɦ] within the minimal prosodic word. Another example is voiceless plosives, which are aspirated initially in the minimal prosodic word but unaspirated otherwise. See Table 1 for examples.

Table 1. Minimal prosodic words. The examples of /h/ are taken from Myrberg & Riad (2015).

mähä	(ˈmäː)$_{\omega min}$(ˌhäː)$_{\omega min}$	’milksop’
koherens	(ko.e.ˈrens)$_{\omega min}$ or (ko.ɦe.ˈrens)$_{\omega min}$	’coherence’
tältöppning	(ˈtʰält)$_{\omega min}$ (ˌöppning)$_{\omega min}$	’tent opening’
kanelte	(ˈkanel)$_{\omega min}$(ˌtʰe)$_{\omega min}$	’cinnamon tea’

The maximal prosodic word ($_{\omega max}$) is the unit which is assigned the accent. For simplex words, the maximal prosodic word is the same as the minimal prosodic word, and the accent is assigned to the stressed syllable. For compounds, which can consist of two or more minimal prosodic words and have both a primary and a secondary stress, only one accent is assigned, and in Central Swedish it is always accent 2. The accent is always assigned to the primary stressed syllable of the maximal prosodic word. Examples are given in Table 2. The acoustic realizations of the accents will be accounted for below.

Table 2. Maximal prosodic words.

tält	^{1}(ˈtält)$_{\omega min=max}$	’tent’
tältöppning	^{2}((ˈtält)$_{\omega min}$ (ˌöppning)$_{\omega min}$)$_{\omega max}$	’tent opening’

The next level in the hierarchy is the phonological phrase. The phonological phrase contains only one and no more than one big accent. Big accent is the term Myrberg and Riad use for what in previous literature has been called focal accent (Bruce, 1977). The reason to call it a big accent instead of focal accent is that there are cases where there is a big accent even though the phonological phrase is not focused (Myrberg, 2010). Initially in an intonational phrase, which is the next level in the hierarchy, for example, there will in most cases be a big accent on what is in traditional North Germanic syntax called the fundament (cf. Roll, 2006). All prosodic words within a phonological phrase that does not contain the big accent get a small accent. Below, the difference between the big and the small accent will be presented. However, they will be called focused (big accent) and unfocused (small accent), since focused and unfocused are the terms used in the game. For unfocused words this means that the stressed syllable will have either a low tone if it is an accent 1 word or a high tone if it is an accent 2 word. Table 3 gives a phonological account of the Central Swedish word accents (Bruce, 1987).

Table 3. The phonology of Central Swedish word accents.

	Accent 1	**Accent 2**
Unfocused	HL*	H*L
Focused	HL*H	H*LH

The different acoustic accent patterns for unfocused words can be seen in Figure 1, and for focused words in Figure 2.

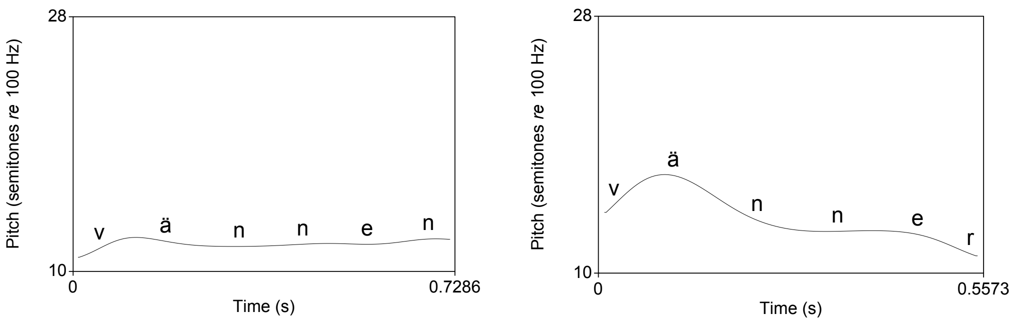


Figure 1. Unfocused Central Swedish word accents on *vännen* ‘the friend’ (accent 1), and *vänner* ‘friends’ (accent 2).

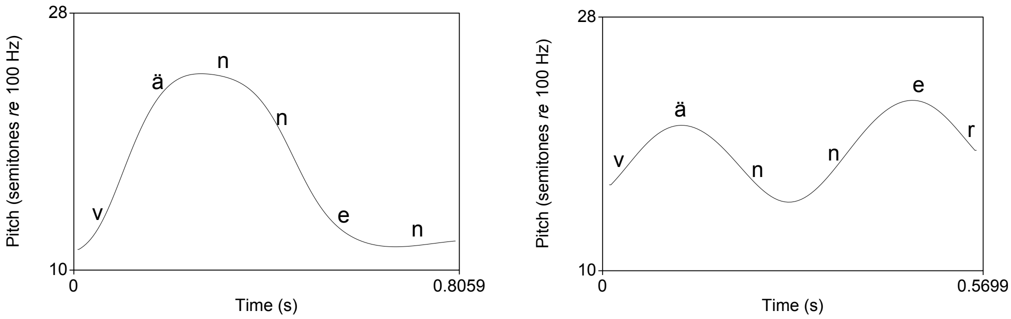


Figure 2. Focused Central Swedish word accents on *vännen* 'the friend' (accent 1), and *vänner* 'friends' (accent 2).

Myrberg and Riad state that the word accents are assigned both at lexical and post-lexical level, resulting in the same surface realization. They argue that accent 2 can be both lexical and post-lexical, but accent 1 is always post-lexical, which means that accent 2 has the status of a phonological tone, while accent 1 is the intonation in lack of this tone. It is worth noting here though that others have argued that both accent 1 and accent 2 have the same phonological status (Bruce, 1977) or that accent 1 is the lexically specified tone, while accent 2 is a result of post-lexical processes (Lahiri et al., 2005).

As mentioned above, the accent is always assigned to the primary stressed syllable of the maximal prosodic word, and following Myrgberg & Riad (2015), accent 2 is specified lexically; either the stem is an accent 2 stem or suffixes induce accent 2. An example where a suffix induces accent 2 on an accent 1 stem is given in Table 4.

Table 4. Accent 2 inducing suffix.

Accent 1 stem	^{1}katt 'cat'
Accent 1 stem + accent 2 suffix	^{2}katt-er 'cat-s'

Accent 2 can also be the result of a post-lexical process, which is what happens when there is a maximal prosodic word containing two or more minimal prosodic words, as seen in Table 2. When there are two stresses in a maximal prosodic word the accent will be accent 2. Accent 1 is always post-lexical, meaning that it is not lexically assigned, and no suffixes induce accent 1 on accent 2 words. However, some atonic prefixes like *be-* and *för-* make accent 2 stems accent 1. The explanation for this is that the prefixes are not part of either the maximal prosodic word, or the minimal. Instead, they are seen as adjoined to the prosodic word, and their special status as atonic prefixes inhibits the accent 2 to be realized. An example of this is given in Table 5.

Table 5. Atonic accent 2-inhibiting prefix.

Accent 2 stem	^{2}ställa 'to put'
Atonic prefix + accent 2 stem	^{1}be-ställa 'to order'

In this study, only simplex words where the minimal and the maximal prosodic word coincide are considered. Only accent 1 stems are used, which are either combined with suffixes that do not induce any accent, or suffixes that induce accent 2. Although the theoretical account states that no suffix induces accent 1, it will be assumed that the tone of the stem is still associated with the suffix that follows. Furthermore, the brain probably uses all kinds of available information for prediction, and thereby accent 1 as well. This will be elaborated upon in the section about previous studies on word accents and neural activity. More information on the stems and suffixes used in this study will be given in the Methods chapter.

When it comes to South Swedish word accents, the rules of assignment are roughly the same, with the exception of compounds, among which there are cases in South Swedish where the compound gets accent 1. The phonetic realization of the word accents in South Swedish differs from their Central Swedish counterparts. One important difference is that where in Central Swedish, the focused accents have an extra H tone compared to unfocused accents, in South Swedish the unfocused and focused accents have the same tonal realization (Bruce & Gårding, 1978). The general difference between the focused and unfocused accents in South Swedish is that the focused accents have a larger fundamental frequency (F0) range and extended duration compared to their unfocused counterparts. A phonological account is given in Table 6.

Table 6. The phonology of South Swedish word accents.

	Accent 1	**Accent 2**
Unfocused	H*L	L*HL
Focused	H*L	L*HL

The acoustic realization of the South Swedish word accents, as produced by a South Swedish L1 speaker is presented in Figures 3 and 4.

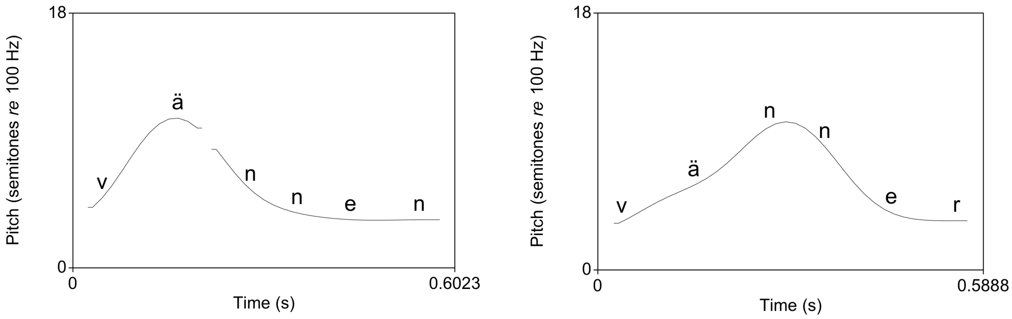


Figure 3. Focused South Swedish word accents on *vännen* 'the friend' (accent 1), and *vänner* 'friends' (accent 2).

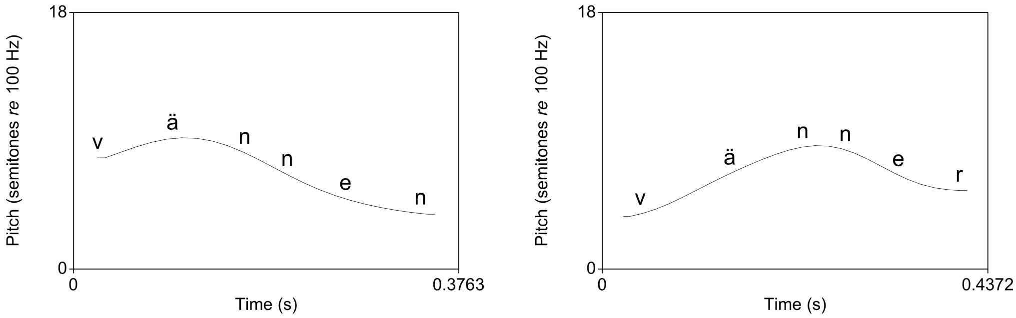


Figure 4. Unfocused South Swedish word accents on *vännen* 'the friend' (accent 1), and *vänner* 'friends' (accent 2).

2.2 Electroencephalography

Electroencephalography (EEG) is a method in which electrodes placed on the scalp pick up electric potential (voltage) fluctuations from the head. The raw data is rather uninformative as regards cognitive functions. Apart from actual electrical activity from the brain it is filled with noise from muscular activity, such as blinks, movements and tension. To extract useful information from this, the event-related potential (ERP) technique is used. ERPs are positive and negative fluctuations triggered by some kind of stimulus. The positivity and negativity have no intrinsic meaning; it is just a matter of where on the scalp the signal is picked up vs. from where it originated. The actual location of the potentials is theoretically impossible to tell from EEG alone. EEG is however very accurate in the time domain, so it is possible to examine these potentials down to milliseconds. To get information about the locality of the neural activity other methods are more useful, such as functional magnetic resonance imaging (fMRI).

To extract ERPs, EEG is recorded from participants in experiments where stimuli of different test conditions are repeated enough times, making it possible to calculate averages of several brain responses to each condition, and by this minimize the signal to noise ratio. Some ERP components have been found a number of times in different experiments. The functions of these different components have been described, and the components have been given names, most often by their polarity, timing and/or locality (Luck, 2005; Roll et al., 2010). Three ERP effects interesting for this study, LAN, PrAN and P600, are presented in section 2.2.1, 2.2.2, and 2.2.3.

Using this method in language-learning experiments, it is possible to get exact information on timing and degree of activity related to a specific language feature. It is also particularly useful in learning experiments, since effects can be found even if there is no change in overt behavior (van Hell & Tokowicz, 2010). A third motivation to use this method is the comparisons that can be made between L1 and L2 language users, e.g. to see which differences there are in terms of the effects found how they are distributed in timing, polarity and amplitude.

2.2.1 LAN

The left anterior negativity (LAN) is an effect reported to appear within 300-600 ms after stimulus onset over, as the name suggests left-lateralized anterior electrode sites, which is mainly elicited by morphosyntactic errors. One example is stimuli where regular rules are applied to irregular words. Weyerts et al. (1997), for instance, found a LAN effect in German L1 speakers for nouns that take the irregular *–n* plural suffix (*Muskel-n* 'muscle-s') when instead the regular plural *–s* was attached (**muskels*). As the name suggests, it is an increased negativity appearing over left anterior electrodes. LAN occurs roughly in the same time window as the N400 component which is a negativity peaking around 400 ms after stimulus over central and parietal electrode sites, but the difference between them is that N400 is more related to semantic processing, whereas LAN is more related to morphosyntax. All words produce an N400, but lexical semantic anomalies elicit enhanced N400 amplitude (Luck, 2005). LAN is often followed by a P600 (see section 2.2.3) in L1 speakers. LAN has been reported to be found for L2 learners as well, especially for highly proficient L2 learners. In a study where German L1 speakers were tested on Italian verb-subject agreement and vice versa, the L2 speakers with lower proficiency showed only a P600, but speakers with higher proficiency showed the biphasic pattern of a LAN followed by a P600 (Rossi et al., 2006). A LAN has also been reported for Swedish L1 speakers in invalid word accent – suffix combinations 200-400 ms after suffix onset, but only in an experiment using pseudowords (Söderström et al. in press).

2.2.2 PrAN

Pre-activation negativity (PrAN) is a slightly left-lateralized negativity proposed to indicate the pre-activation of upcoming linguistic material (Söderström et al., submitted). It correlates with type frequency, where a segment with fewer types yields a larger PrAN than a segment with more types. This is found for L1 speakers of Swedish at 136-280 ms following word accent onset where Accent 1 yields increased negativity. This is because, as stated in section 2.1, accent 1 is associated with a small, well-defined set of endings, whereas accent 2 can have almost infinite continuations due to the compounding rule (Roll et al, 2015; Roll, 2015; Söderström et al., submitted; in press). Therefore, accent 1 can be argued to more strongly activate its associated suffixes. There has been one known study where PrAN has been investigated in L2 learners. Gosselke Berthelsen et al. (in preparation) tested low to intermediate level L2 Swedish speakers with German L1 on Swedish word accents, but did not find any PrAN in the neural activity.

2.2.3 P600

The P600 is a positive fluctuation peaking around 600 ms after stimulus onset, for stimuli containing syntactic violations. It is proposed to be an indication of reanalysis and repair of incorrect or effortful forms. It is normally found at posterior electrode sites. For Swedish word accents, a P600 is found between 400 and 900 ms after onset of the suffix in invalid accent - suffix combinations (Roll et al., 2010; 2013; 2015; Söderström et al., in press). P600 has also been found for L2 learners, as stated above, but one notable difference is that it is reported to appear slightly later for L2 learners than for L1 learners (Luck, 2005; van Hell & Tokowicz, 2010).

2.3 Word accents and L1 neural activity

As mentioned in section 2.1, suffixes induce word accents occurring before the suffix itself, and this has led to the hypothesis that Swedish L1 speakers might be able to predict the suffix or word form from just hearing the accent. Some work has been done to corroborate this hypothesis, as briefly touched upon, and different psycho- and neurolinguistic methods have been used. These methods include reaction time studies (Söderström et al., 2012), EEG (Roll et al., 2010; 2013; 2015; Roll, 2015; Söderström et al., in press), and fMRI (Roll et al., 2015). Both accent 1 and accent 2 are used as predictors for suffixes, but accent 1 seems to be a stronger predictor. This is shown by response times and ERPs. In response time tests, where participants have listened to sentences with target words containing valid and invalid accent - suffix combinations, and have been asked about the meaning of the target words, accent 1 has yielded significantly shorter response times for valid suffixes than for invalid ones. Table 7 gives examples of valid and invalid accent-suffix combinations.

Table 7. Valid and invalid accent - suffix combinations.

	Valid suffix	**Invalid suffix**
Accent 1	^{1}fisk-en 'fish-the'	*^{1}fisk-ar
Accent 2	^{2}fisk-ar 'fish-pl'	*^{2}fisk-en

In EEG data, accent 1 gives a stronger pre-activation negativity (PrAN), which correlates with shorter response times. PrAN, and thereby the differentiation between accent 1 and 2, is visible 136 ms after accent onset and is left lateralized (Söderström et al., submitted; Roll et al., 2015). It has also been shown that it is the phonology of the tones, and not the F0 in itself that yields these results, since the same pattern for accent 1 and accent 2 was found for South Swedish speakers as well. In South Swedish, accent 1 is a high tone, and accent 2 is a low tone, i.e. the opposite of Central Swedish (Roll, 2015). Another corroboration of the hypothesis is that this effect is not found when L1 speakers listen to non-linguistic equivalents to the intonation of the accents. This is shown in an experiment where participants listened to hummed versions of the accents (Roll et al., 2013). In Söderström et al. (in press), the stimuli consisted of pseudowords with possible Swedish phonotactic combinations. This goes to show that the accent can yield a PrAN without lexical content.

Söderström et al. (submitted) showed that the PrAN effect correlates with type frequency. Thus, word-initial fragments with fewer possible continuations gave rise to a larger PrAN. This indicates that the PrAN might be an index of certainty about the continuation of a word. Accent 1 has a small, well-defined set of endings associated to it and thereby fewer types. Accent 2 is the accent induced by compounding, which is a very frequent and productive morphological process in Swedish, and thereby has more types. Therefore, the PrAN is increased for accent 1.

A left-anterior negativity (LAN) (Söderström et al., in press) or a broadly distributed negativity (Roll, 2015) followed by a P600 (Söderström et al., in press; Roll, 2015; Roll et al., 2013; 2015) have also been found from suffix onset when participants heard accents with their invalid suffixes and were asked to press one button if the word meant “one thing” and another button if the word meant “many things” (i.e. had the definite singular suffix *-en* or the indefinite plural suffix *-ar*) (Roll et al., 2013). Along with the P600, longer response times for invalid suffixes are found (Söderström et al., 2012; Roll et al., 2013). It is worth noting that a LAN effect has only been found in a study using pseudowords as stimuli.

In an fMRI study on suffix activation by word accents in L1 speakers of Swedish, it was found that word accents activate Brodmann Area 22 in the left superior temporal gyrus. This is the same area found for e.g. Thai tones, which corroborates the notion that the Swedish word accents are indeed lexical and not a part of sentence intonation, which has been argued to rely more on the right side of the brain for its processing. After identification, suffix activation is thought to be modulated in the left inferior frontal gyrus, Brodmann area 47. Affix processing has been reported to activate that area in other studies. Accent 1 gives increased activation, which is consistent with the notion that it is a better predictor than accent 2. About 120 milliseconds after this, morphological processing is thought to increase involvement of the inferior frontal gyrus (Roll et al., 2015).

2.4 Second language acquisition of word accents and tone

There is not much focus on the training of word accents in courses on Swedish as a second language. They are not seen as very important, since the number of minimal pairs is low, and Swedish L1 speakers in general have a low meta-understanding of them. There are also dialects of Swedish with no word accent distinction at all, such as the varieties of Swedish spoken in Finland, which are interintelligible with other varieties of Swedish, which shows that word accents are not crucial for making oneself understood. However, if word accents are good predictors for different suffixes for L1 speakers of Swedish, they might also serve to facilitate processing in L2 speakers. Thus, it may be a good idea to include teaching of them in class. Therefore, one aim of this study is to investigate if and how L2 learners of Swedish can be trained to use word accents as grammatical predictors.

The field of research concerning second language acquisition (SLA) of Swedish word accents is still small, and only one previous study with a neurolinguistic approach (Gosselke Berthelsen et al., in preparation) is found in the literature. Hed (2014) tested whether Somali L1 speakers would perceive and produce Swedish word accents. Somali is also a language with word accents, and it has been proposed that L1 speakers of word accent languages are facilitated in acquiring word accents in a tonal L2. The study consisted of a production test, where the participants read a text with target words that were verbs either in present tense *springer* ’runs’ with accent 1, or the infinite form *springa* ’to run’ with accent 2. The study could show that the speakers did the distinction in production, but in a discrimination test, they were not better than chance in perceiving the difference. The participants did not have any formal education about the Swedish word accents and the grammatical connection, which indicates that word accent production can be acquired without any production training. Other studies on SLA of Swedish word accents have proposed that speakers of non tonal languages (German, Farsi) did not produce or perceive the word accent distinction, (Tronnier & Zetterholm, 2013; Kaiser, 2011), and in Tronnier and Zetterholm’s study it was also proposed that speakers of Vietnamese, a tonal language, did not produce the word accent difference. Somali L1 speakers were also tested by Tronnier & Zetterholm (2013), and they proposed that these speakers did produce the word accent difference. However, in a perception test where Swedish L1 speakers evaluated the difference between the word accents produced by the Somali L1 speakers, the Swedish speakers did not perceive a difference (Tronnier & Zetterholm, 2014).

Van Dommelen & Husby (2009) studied the acquisition of Norwegian word accents, which are similar to Swedish word accents, by Mandarin and German L1 speakers. They posed the question whether the L1 would affect acquisition since

Mandarin is a tonal language and German is not. They conducted identification and discrimination experiments and concluded that the Mandarin L1 speakers perceived word accents with a higher accuracy than the German L1 speakers. They also posed the question if training would improve word accent acquisition. Therefore, half of the participants got a one-hour lecture on the word accents, and after that all participants underwent the tests again. Both participants with and without training improved their results in the discrimination test, but participants with training improved more. Improvement in participants that did not take the lecture was explained by familiarization effects of the test situation. For the identification test, participants without training did not improve their results, while participants with training did; however, the difference was only statistically significant for the Mandarin L1 speakers, and not for the German L1 speakers. This study only made use of behavioral measures. It would be interesting to see what the results would look like in an additional ERP study. It can also be noted that the training in this case was more theoretical than practical.

Two studies of special interest for this thesis are the studies by Schremm et al. (2016) and Gosselke Berthelsen et al. (in preparation), as briefly described in the Introduction chapter. Both studies used similar methods to the ones described for L1 speakers in section 2.3. Gosselke Berthelsen et al. tested German L1 speakers of low to intermediate level of Swedish in an ERP study and found no PrAN effect, indicating that these L2 learners had not acquired the predictive function of word accents. Schremm et al. tested non-tonal L2 speakers at a higher level in a behavioral study, and found that they had longer response times when the word accent did not match the suffix. This indicates that L2 learners implicitly do learn to use word accents as predictors in a similar manner to L1 speakers.

Where there is a lack of research on word accent acquisition, more work has been done on tone acquisition, with languages such as Mandarin and Thai in focus. In these languages, every syllable is assigned a tone, and the tone changes the semantics of the syllable, whereas Swedish has one accent per word that is, as stated in section 2.1, associated with morphological affixes

Several studies have shown that there is an effect of training on tone perception. Wang et al. (1999) trained one group of American English L1 speakers in Mandarin tones, and contrasted them with a control group of American English L1 speakers who did not participate in any training. All were beginner L2 learners of Mandarin. They did a pre-test and a post-test where the participants were asked to identify what tone a word in isolation had. In between the tests, the training group participated in 8 training sessions of 40 minutes each during two weeks, where they trained tone identification. The participants that received training had a 21 % increase in accuracy, and they also increased their results with novel stimuli and voices. The effect of the training was still present after six months. The control group did not increase their accuracy. Wayland & Guion (2004) conclude that after 5 days of training with 30-minute training sessions, naïve L1 speakers of Mandarin and Taiwanese improved their perception of Thai tones in a discrimination test, while naïve L1 speakers of English did not improve. However, all improved their results in an identification test. The training consisted of a discrimination exercise where the participants heard three different Thai syllables where either none or one of the instances had a different tone compared to the others. The participants were asked which of the syllables differed from the others. Francis et al. (2008) investigated second language perception of Cantonese by English and Mandarin L1 speakers. They tested this with a pre- and post-test consisting of both identification and difference rating. They used 7 different carrier sentences, with 6 different syllables and 6 tones. Between the pre- and post-test the participants trained for 14 days, and within these 14 days there were 10 training days. The training was a game where the participants listened to a target word embedded in a carrier sentence, and they were asked to press a button with the correct tone. The forced choice alternatives were presented with a Chinese character, a Pinyin gloss, the number of the tone, and a graphical picture of the F0 contour. In the pre-test, the two groups did not differ from each other. In the post-test, results improved in both groups, but they made different kinds of mistakes. The authors interpret this as an L1 effect.

Since there seems to be a difference in how L1 learners with and without tone in their first language acquire tone, only L1 learners with non-tonal L1s are included in the present study.

There is also evidence that perception training without any production training enhances production. Wang et al. (2003a) showed this in a study where participants were recorded when they read a list of isolated words before and after 8 sessions of 40 minutes perception training, the same as in Wang (1999). Mandarin L1 speakers evaluated the tones by listening to the recordings and choosing which tone it was, or if there was none at all. The authors concluded that the participants with training improved their production, while the ones without training did not, which implies that perception training enhances production.

Most of these studies only provide behavioral data, but Kaan et al. (2007) also concluded that 2 days of 30 minute training on three different Thai tones with a passive oddball experiment changed the ERPs for naïve Mandarin and English L1 speakers. The tones in the experiment elicited different ERPs. Low-falling elicited mismatch negativity (MMN) in both groups before and after training, High-rising got a smaller MMN, but it increased for the English speakers after training. High-rising caused a late negativity, which was reduced after training for the Mandarin speakers. Lu et al. (2015) investigated the effect of perception versus joint perception and production training on learning of tones. English L1 speakers participated in a pre- and post-test with a passive oddball experiment. ERPs were recorded while participants were watching a silent movie and listening to stimuli, which were presented in headphones with one of the tones as a standard and the other two as oddballs. The stimuli were 8 Mandarin-like syllables with tones similar to Mandarin tones 1, 2 and 4. There was also a forced choice AX-test, where two tones were played in a row and the participants were asked if the second one was similar or dissimilar to the first one, to obtain behavioral data. After this, half of the participants performed a perception exercise for one hour where they listened to the syllables with the different tones and uttered a completely different word afterwards. The other half performed a joint perception and production exercise, where they listened to the same stimuli but repeated the syllable with its tone afterwards. Results showed that accuracy increased in all participants, and the ones who underwent the joint perception and production exercise did not improve more than the others. For the ERPs, there was increased MMN pre training, which decreased post training for both groups. The explanation given was that pre training the participants were more sensitive to the F0 onset difference, whereas they paid more attention to F0 direction post training.

Positron Emission Tomography (PET) studies have shown that in speakers of non-tonal languages, word tone processing is modulated mainly by areas in the right hemisphere of the brain, showing patterns overlapping with non-linguistic processing of pitch (Klein et al., 2001). Wang et al., (2003b) did an fMRI study on low level American L1 speakers learning Mandarin before and after a two-week training period in identifying the Mandarin tones. They could show that after the training period, activation increased in Brodmann area 22 in the left superior temporal gyrus, and new areas were activated as well, such as Brodmann area 42 in the left superior temporal gyrus. These areas are the same as where L1 speakers of tonal languages process tone, as mentioned in section 2.3.

Something that has not been investigated in the L2 studies presented here, but will be in the present study, is the association between tones and grammar/lexicon, instead of the focus on differentiating between tones. The previous studies on L2 acquisition pose the questions *"which tone is this?"* or *"are the tones different?"* rather than *"what does this tone imply?"* as will be done in this study. It can be argued that this gives this study more ecological validity, since the training focuses more on how it is known that L1 speakers use the accents, and not a meta understanding of them. Swedish with its morphological assignment of word accents provides an excellent opportunity to test how L2 speakers can use the intonational information in a more natural way.

2.5 Research questions and hypotheses

With the theoretical background presented above, this section gives a more in-depth discussion of the research questions and hypotheses presented in the introduction chapter. The first question is *how production, perception and neural activity change during and after the two-week training period.* This is the most basic question of the study. Gosselke Berthelsen et al. (in preparation) did not find any PrAN effect for early L2 learners of Swedish, but Schremm et al. (2016) found that learners with higher proficiency did make use of the predictive nature of word accents. It is in the gap between these studies that this thesis takes its point of departure. Hence, the question is whether two weeks of training is sufficient to see any changes in production and perception. According to previous literature, it should be enough to see change in perception behavior (Van Dommelen & Husby, 2009; Wang, 1999; Wayland & Guion, 2004; Francis et al., 2008), ERPs (Kaan et al., 2007) and production (Wang et al., 2003a). Of interest is also the notion that ERP effects can be found even if there is no effect in behavioral measures (van Hell & Tokowicz, 2010).

The second question is *whether the production, perception and neural activity will be more L1-like after training.* Support for the hypothesis that it will be is found in Wang (2003b), who reported that the neural response to tone becomes more left lateralized, and thereby more L1-like, for L2 learners of Mandarin Chinese as their proficiency increases. The source of the electrical activity is, as stated in section 2.2, theoretically impossible to track with EEG, but temporal and topographical distribution of effects might give a hint for future studies. Schremm et al. (2016) also found that L2 learners with higher proficiency were L1-like in their behavior, and this came from implicit learning, with no special training in word accents at all. Comparisons will also be made with studies on Swedish L1 speakers (Söderström et al., 2012; submitted; in press; Roll et al., 2010; 2013; 2015; Roll, 2015) to look for similarities and L2-specific processes.

The research questions and the previous literature lead to three main hypotheses. The first one is that *training will lead to increased predictive use of word accents.* The second one is that *perception training will lead to enhanced production*. Thirdly, it is hypothesized that *production, perception and neural activity will be more L1-like after training*.

These hypotheses lead to a number of predictions to be tested in the study. First of all, considering the ERPs, two onsets will be used: word accent onset and suffix onset (more on this in the Methods section). For the accent ERPs, one prediction is that there will be increased negativity (PrAN) post training. This will indicate that L2 learners can learn to use word accents in a similar predictive fashion as L1 speakers do (Söderström 2015 et al., submitted). Another prediction is that the accent processing will be more left-lateralized post training. This is, again, something that cannot be directly ascertained from EEG data, but there might be a hint about it in the topographical distribution of ERP effects.

As for the suffix ERPs, two effects are predicted not to be found pre training, but post training. The first is a negativity, perhaps with a left anterior distribution (LAN), which is predicted to be found for invalid accent - suffix combinations post training. A broadly distributed negativity was found for L1 speakers in Roll (2015), and a LAN was found in Söderström et al. (in press) who used pseudowords as stimuli. The second one is a P600, which would indicate a reprocessing of the invalid accent - suffix stimuli post training. These ERP effects are the ones that can be found in L1 speakers' neural responses to these stimuli (Roll, 2015; Söderström et al., in press) and it is hypothesized that L2 learners will get closer to L1 processing of the word accents.

For the behavioral data of the training it is expected that response times will decrease and accuracy will increase during the duration of the game, indicating that the participants will learn the connection and become faster and give more accurate answers.

In the pre- and post-test, some behavioral differences are predicted to be found before and after training. For one, post-training perception results are expected to have overall shorter response times. Another prediction is that pre training, there will not be any effect for invalid accent - suffix stimuli but post training there will be longer response times for these. If this is the case the results will be more similar to Swedish L1 speakers (Söderström et al, 2012) and L2 speakers with higher proficiency (Schremm et al, 2016), who show increased response time for invalid stimuli.

The last part of this section considers production. The hypothesis is that production will be enhanced post training. Wang (2003a) states that production can be enhanced by perception training, and according to this study, two weeks should be enough to detect effects. Hed (2014) also concludes that production of word accents can be acquired without special training, at least for L1 speakers of another word accent language.

3 Methods

3.1 EEG

The EEG-data was recorded with a BRAINAMP MR PLUS Amplifier in Brainvision Recorder, both from BrainProducts. The cap used was a 32 Channel Standard Braincap-MR from Easycap. The electrode set-up is given in Appendix 1. The data was recorded at a sampling rate of 250 Hz. A high-pass filter with cutoff frequency at 0.1 Hz and a low-pass filter with 70Hz cutoff were applied online. Impedances were kept below 5kΩ and FCz was used as an online reference.

For the analysis, BrainVision Analyzer was used. The data was re-referenced, using an average of the mastoids (TP9 and TP10). An offline low-pass filter with a cutoff frequency of 30Hz was used. Vertical and horizontal eye movements were compensated for using ICA. Epochs were extracted, starting from accent and suffix onset to 1000 ms post onset. A window of 200 ms before this was used as a baseline. Epochs were rejected where voltage step exceeded 50 µV/ms. Maximal allowed difference of values in intervals was 200 µV and lowest allowed activity in intervals was 0.5 µV. On these grounds <1% of the epochs were rejected.

For statistical analysis, different averages were extracted:

- 136-280 ms after accent onset. This is where it is known that Swedish L1 speakers process and differentiate between the word accents.
- 225-300 ms after suffix onset, where visual inspection suggested an increased negativity (LAN) post game.
- 325-425 ms after suffix onset. This is where the suffixes seemed to have an effect independently of word accent.
- 400-600 ms after suffix onset, a time window in which visual inspection suggested a possible P600 effect.

For the word accent analysis, repeated measure ANOVAs were performed with factors session (1 and 2), accent (1 and 2), and region of interest (ROI). The ROIs used were Left Anterior (LA) consisting of F3 and F7, Right Anterior (RA) consisting of F4 and F8, Left Central (LC) consisting of C3 and T7, Right Central (RC) consisting of C4 and T8, Left Posterior (LP) consisting of P3 and P7 and Right Posterior (RP) consisting of P4 and P8.

For the suffix analysis, both accent and suffix were tested as factors along with session, validity and ROI.

3.2 Pre- and post-test

The pre- and post-test experiment was a response time test, previously used in a combined EEG and fMRI study on Swedish L1 speakers (Roll et al., 2015). Stimulus sentences were played in loudspeakers. The target word had either accent 1 or accent 2 on the stem, and the singular definite suffix *-en* or the plural indefinite suffix *–ar*. An example sentence is presented in Table 8.

Table 8. Example sentences in the pre- and post-test.

Rut fick fisk-en till lunch.	Rut fick fisk-ar till lunch.
'Rut got fish-the for lunch.'	'Rut got fish-pl for lunch.'

The participants were asked if the person talked about in the sentence got one or many things. They answered by pressing one of two buttons on a response box, using their left and right index fingers.

The stimuli had four conditions, presented in Table 9.

Table 9. Stimuli conditions in the pre- and post-test.

	Accent 1	**Accent 2**
Singular suffix *–en*	accent 1 + *-en*	accent2 + -**en*
Plural suffix *–ar*	accent 1 + -**ar*	accent 2 + *-ar*

Thus, half of the material consisted of valid word accent - suffix combinations and half of the material consisted of invalid word accent - suffix combinations. There were 30 sentences per condition, making it 120 sentences in total. The test had two blocks of equal length, and in the different blocks, either right or left index finger was *singular*. Half of the participants started with left index finger as *singular* and half of the participants started with right to control for bias effects. The interstimulus intervals varied between 4 to 8 seconds. The stimulus order was pseudo-randomized and the order of the buttons was counter-balanced over the blocks. Instructions were given in both Swedish and English. The test took approximately 20 minutes, with a break of self-imposed length in the middle.

In experiments with Swedish L1 speakers, subjects' responses have been slower for the invalid suffix - word accent combinations. The interpretation of this is that the subjects use word accents as predictors for the suffixes and have to reprocess the invalid word structure if an invalid suffix follows the accent, and hence take longer to respond. This is also visible in the EEG as a P600 (Roll, 2015; Roll et al., 2010; 2013; Söderström et al. in press). The idea in this experiment is to see to what extent the L2 learners use the accents as predictors before and after training.

In the pre- and posttests, the target word was always in unfocused position, with the focus on the last word of the sentence. In the sentence presented in Table 8, the focus was on *lunch*. The word before the target word was always *fick* ’got’, ending with a voiceless stop, and the word after the target word was always *till* ’to’, starting with a voiceless stop. This was to make cutting and concatenating recorded sentences easier. For the same reason, the monosyllabic stem always ended in a voiceless stop. The sentences were recorded in an anechoic chamber by a male Central Swedish speaker and then cut out so the target words were combined with different carrier sentences. Half of the accent 1 sentences had target sentences from accent 2 recordings and vice versa.

The test was conducted using E-prime software on a PC with answers given by response box button press. The stimuli were presented in speakers with sound at a comfortable level.

For analysis, response times under 0 ms and over 2000 ms were excluded.

For statistics, repeated measure ANOVAs were made with factors session (pre and post), accent (1 and 2) and suffix (valid, invalid). Accuracy was also tested pre and post training.

3.3 LMG – The Language Melody Game

The Language Melody Game (LMG) is a game in development, and it is to be used as a smart phone application for training the connection between word accents and grammatical suffixes as described in section 2.1. For this study a website version, available for the participants online, was used, developed by Andrea Schremm.

The idea of the game is that the player gets to hear a sentence up to the stressed syllable of a target word. After hearing the word accent of the syllable, the player gets to choose between two different suffixes by clicking one of two buttons. The sentences within the game looked like in Figure 5.

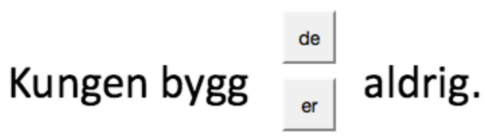


Figure 5. Example of sentence and layout in the Language Melody Game. 'The king never build-s/t.'

There were 3 levels in the game. The first one contained target words in focused position, the second one words in unfocused position, and the third one was mixed. Within the first 2 levels there were 3 sublevels: nouns, verbs and mixed parts of speech. In the third level everything was mixed. Each sublevel contained 100 pairs of words with accent 1 + suffix 1 and accent 2 + suffix 2. In total, 800 sentences were recorded. However, one focused noun sentence pair and two unfocused noun sentence pairs were excluded due to artifacts in the recordings. Thus, the total number of sentences used was 796. The sublevels were also divided into rounds with 18-20 sentences in each. Every round contained equally many accent 1 and accent 2 sentences. The structure of the game is shown in Figure 6.

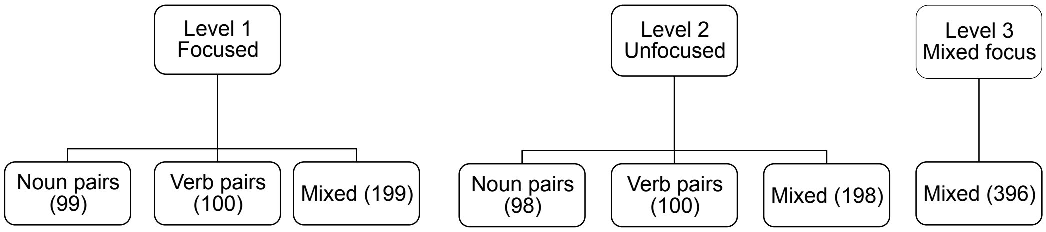


Figure 6. Game structure in the Language Melody Game.

All stems were monosyllabic. For the nouns, they were either in definite singular form, with accent 1 and the definite singular *–en* suffix; or in indefinite plural form with accent 2 and the indefinite plural allomorphic suffixes *–ar* or *–er*. Twenty of the most frequent nouns in the Stockholm-Umeå Corpus (SUC) (Ejerhed et al., 1992) were used. There were 5 different carrier sentences. For the verbs, the accent 1 forms were in present tense with the suffix *–er*. The accent 2 forms were in preterite tense with the allomorphic suffixes *–te*, *–de* or *–e*. The 45 most frequent verbs in SUC were used in 8 different carrier sentences. Examples of sentences are given in Table 10 and a full list of the words used is presented in Appendix 2.

As mentioned in section 2.4, one notable difference from other studies on L2 perception of tone is that this study did not ask for *"is this tone x or y?"* but rather involved a semantic task related to the meaning associated with the tone - suffix combination.

The stimulus sentences were recorded by a female speaker of Central Swedish in an anechoic chamber. The speaker was asked to read the sentences from a paper, each accompanied with a context question to ensure that the focus would be placed at the right word. Examples are given in Table 10.

Table 10. Examples of sentences in The Language Melody Game

Sentences in focused position	
Vad läser vi om i en bok? ’What do we read about in a book?’	Vi läser om tid-en/er i en bok. ’We read about time-the/s in a book.
’Vad gjorde kungen aldrig? ’What did the king never do?’	Kungen ring-er/de aldrig. ’The king never call-s/ed.’
Sentences in unfocused position	
Var läser vi om tiden/er? ’Where do we read about time-the/s?’	Vi läser om tid-en/er i en bok. ’We read about time-the/s in a book.’
När ringer/de kungen? ’When does/did the king call?’	Kungen ring-er/de aldrig. ’The king never call-s/ed.’

The sentences were randomized to control for other possible intonational cues such as left edge boundary tones/initiality accents (Roll, 2006; Myrberg, 2010). The sentences were read twice to ensure accuracy. Afterwards, they were controlled by the author and other phonetically trained persons and excluded if there were artifacts. As mentioned above, 3 pairs of sentences were excluded. The sentences were then cut after the stem syllable, at zero crossings, using Praat (Boersma & Weenink, 2014). The amplitude of the sentences was normalized to an average volume of 70 dB.

In the game the whole sentence was presented in text, with the two alternative suffixes, and the first part of the sentence was played. See Figure 5 for an example sentence. The subject then clicked on the alternative they thought to be the correct one. The alternatives varied so that as many singular/present were the upper alternative as the plural/preterite. When the player pressed one of the alternatives, immediate feedback was given; the whole correct sentence was shown and if the answer was correct it turned green, otherwise it stayed black. The correct ending of the sentence was also played.

The participants were asked to play the game for 10 days during a period of 14 days, and for each play day they played for at least 15 minutes and maximum one hour. They played at home and could choose which days they played and if they wanted to use speakers or headphones. The amount of time spent on the game as well as levels reached and accuracy was logged.

The participants moved to the next sublevel when they reached 80 % correct answers during 1 round, or after a fixed amount of time, to make sure that all participants moved through every level. These times were 1 play-day for level 1 and 2 play-days for level 2.

For analysis some measures were taken from the game data:

- Amount of time spent playing the game, called playtime, to see if there was any correlation between playtime and changes in the EEG pre and post training.
- An average of the last five rounds for each participant, called final accuracy, also this to correlate with EEG data.
- An average of response time for correct responses per stage. Times over 5000 ms and negative times were excluded, to see development during the game.
- An average of accuracy per stage. This was also used to see development during the game.

3.4 Production

A production test was also conducted pre and post training. The participants were asked to read 20 sentences as answers to context questions. The context questions were given to try to elicit focus on the same place in all sentences, but only the answer was recorded. The structure of the sentences was the same as in the game, where they would contain either accent 1 or accent 2 target words presented in a carrier sentences. 10 of the sentences had target nouns in either singular or plural, with the valid accent: accent 1 for singular definite and accent 2 for plural indefinite. There were two carrier sentences, one for nouns and one for verbs. 10 of the sentences had target verbs in either present or past tense, with valid accents: accent 1 for present tense and accent 2 for past tense. Half of the words were present in the game and half were not, to control for rule generalization. Examples are given in Table 11 and the full lists are given in Appendix 3 and 4.

Table 11. Example sentences from the production test.

Var såg Anna dagar? ’Where did Anna see days?’	Anna såg dagar på TV. ’Anna saw days on TV.’
Vad glömmer Tom? ’What does Tom forget?’	Tom glömmer saker. ’Tom forgets things.

They were recorded with a portable recorder (TASCAM DR-07) in the same room as the EEG recording was conducted. The experiment leader left the room while the recording was made to decrease nervousness and discomfort for the participant. The recordings were cut and analyzed in Praat (Boersma & Weenink, 2014), where the author audiovisually evaluated the intonation of the target word. The target word was coded as either focused accent 1 (1), focused accent 2 (2), unfocused accent 1 (3) or unfocused accent 2 (4) if the F0 curve sounded like or resembled the Central Swedish accents. It was also coded as unsure (5) if it did not resemble any of the focused and unfocused accents. An extra code (6) was also added since some of the words more closely resembled South Swedish accents. The author is an L1 speaker of Central Swedish and also trained in phonetics, hence, her audiovisual inspection can be considered a fairly accurate one. Another trained phonetician, although with South Swedish L1, also annotated 30 % of the material to check for inter-rater agreement, using Cohen's kappa (Cohen, 1960). Figure 7 shows an example of the annotated text grid in the Praat window.

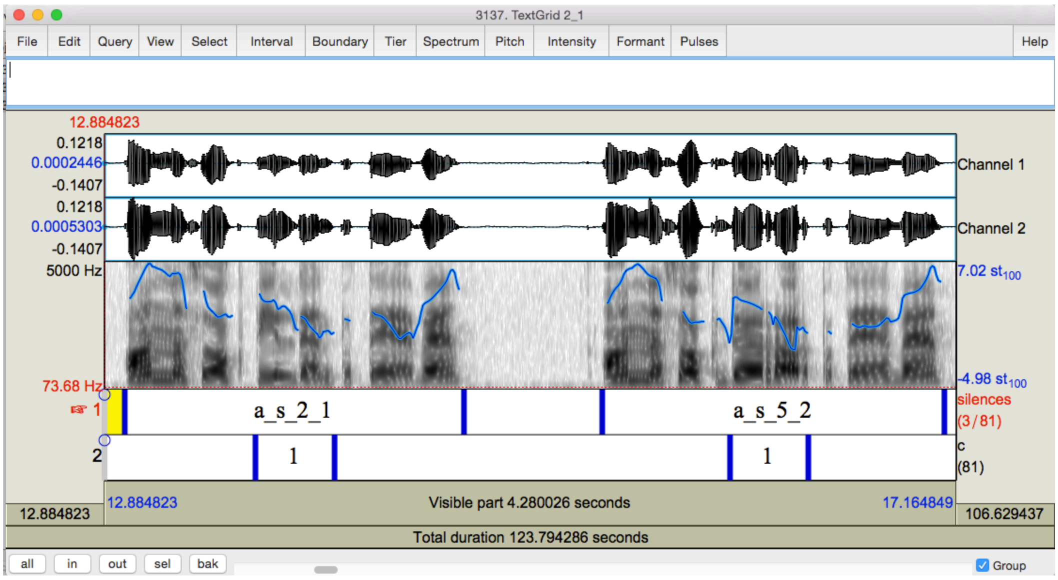


Figure 7. Annotation example for the production test, conducted in Praat. Sentence a_s_2_1 (a = pretest, s = *substantiv* 'noun', 2 = code of the target word, 1 = accent 1 target word) *Anna såg bilden på TV* 'Anna saw the picture on TV', where *bilden* is the target word. Sentence a_s_5_2 (a = pretest, s = *substantiv* 'noun', 2 = code of the target word, 2 = accent 2 target word) is *Anna såg dörrar på TV* 'Anna saw doors on TV', where *dörrar* is the target word.

Repeated measures ANOVAs were made on the results with factors session (pre and post), accent (1 and 2) and game presence (yes or no). Game presence was considered as a factor to see if the participants acquired the word with the accent as a unit or if they acquired a rule that generalized over novel words, not present in the game. Accuracy post training and the difference in accuracy pre- and post training were also correlated with ERP data to see if participants who performed with higher accuracy differed in their neural response compared to participants with lower accuracy; with playtime from the game, to see if participants who trained more were more accurate in production; and with final accuracy from the game to see if the participants with higher accuracy in perception also showed higher accuracy in production.

3.5 Participants

Nineteen speakers of non-tonal languages participated in the study. They were between 20 to 26 years old with an average of 23 years old (*SD=1.8*). They self-reported their knowledge of Swedish, showing that they had studied Swedish on an average of 22.2 (*SD=24.6*) months, and were between new beginners up to level B2. They had been living in Sweden for an average of 5.8 (*SD=4.7*) months. All of them studied or were about to start studying Swedish at the courses for international students at the university or at SFI (Swedish for Immigrants). They were all students at bachelor or Masters level at Lund University. Some had studied Swedish before coming to Sweden and some started their experience of Swedish in Sweden. A more extensive summary of the participants is given in Appendix 5.

4 Results

4.1 EEG data

4.1.1 Word Accent ERPs

There was a main effect for session at 136-280 ms following F0 onset, $F(1,18) = 5.929$, $p = 0.026$, which showed increased negativity in the post-test. The mean values pre and post are given in Table 12.

Table 12. Mean values pre- and post test 136-280 ms after accent onset.

	Mean (*SD*) in µV
Pre-test	-0.327 (*0.141*)
Post-test	-0.581 (*0.112*)

There was also a main effect for accent, $F(1,18) = 8.332$, $p = 0.010$, where accent 1 yielded more negativity both in the pre- and the post-test. The mean values for the accents pre and post game at all ROIs are given in Table 13.

Table 13. Mean values for each accent pre and post training, 136-280 ms after accent onset.

		Mean (*SD*) in µV
Pre-test	**Accent 1**	-0.441 (*0.162*)
	Accent 2	-0.213 (*0.138*)
Post-test	**Accent 1**	-0.684 (*0.112*)
	Accent 2	-0.447 (*0.131*)

Session further interacted with antpost, showing that negativity increased anteriorly, $F(1,18) = 12.057$, $p = 0.003$, and centrally, $F(1,18) = 9.930$, $p = 0.006$.

Figure 8 shows the increased negativity on RA, where the change was clearly visible.

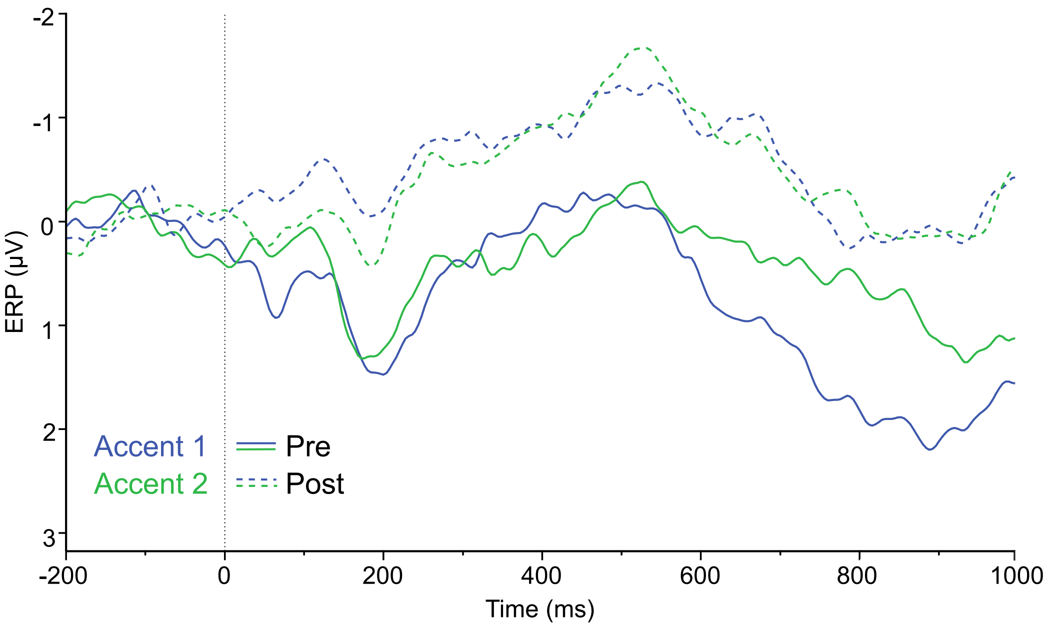


Figure 8. Right Anterior electrode site. 0 = accent onset. Blue line = Accent 1. Green line = Accent 2. Solid line = Pre training. Broken line = Post training.

4.1.2 Suffix ERPs

At 225-300 ms after suffix onset, there was a suffix main effect where the singular suffix yielded more negativity than the plural suffix both pre and post training $F(1,18) = 7.567$, $p = 0.013$. The values pre- and post for all ROIs are given in Table 14.

Table 14. Mean values for each suffix 225-300 ms after suffix onset.

		Mean (*SD*) in μV
Pre-test	***-en***	0.732 (*0.166*)
	-ar	0.874 (*0.193*)
Post-test	***-en***	0.462 (*0.195*)
	-ar	0.730 (*0.194*)

There was also a session × validity × ROI interaction, $F(5,90) = 3.353$, $p = 0.008$. From follow up ANOVAs on each ROI, increased negativity for invalid suffixes was found in the LA, indicated by a session × validity interaction, $F(1,18) = 6.805$, $p = 0.018$. This can be interpreted as LAN for invalidity. The effect can be seen in Figure 9.

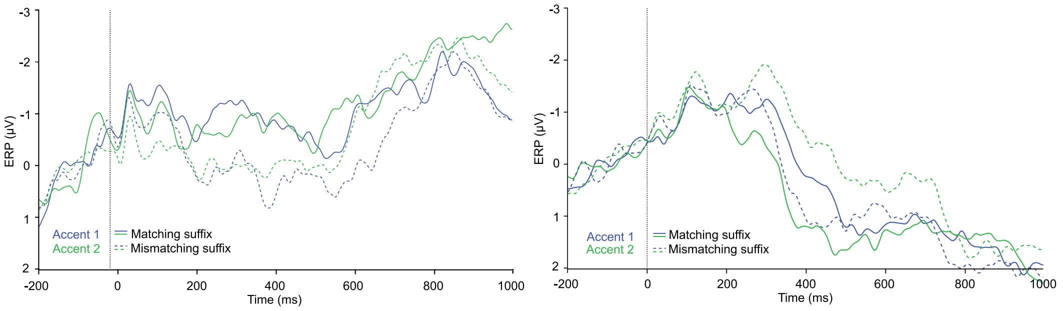


Figure 9. Left Anterior electrode site. 0 = suffix onset. The left one shows pre training and the right one shows post training. Blue line = Accent 1. Green line = Accent 2. Solid line = Matching suffix. Broken line = Mismatching suffix.

325-425 ms after suffix onset the suffix main effect was consistent, the singular suffix was still more negative $F(1,18) = 18.853$, $p = 0.001$. The effect was present both pre, $F(1,18) = 6.472$, $p = 0.018$, and post training, $F(1,18) = 18.128$, $p > 0.001$. The values for all ROIs are given in Table 15.

Table 15. Mean values for each suffix 325-425 ms after suffix onset.

		Mean (*SD*) in µV
Pre-test	***-en***	0.987 (*0.266*)
	-ar	1.372 (*0.240*)
Post-test	***-en***	1.071 (*0.252*)
	-ar	1.976 (*0.297*)

Furthermore, there was a session × suffix interaction, $F(1,18) = 5.885$, $p = 0.026$. There was no session effect for the singular suffix but the plural suffix was more positive in the post-test, $F(1,18) = 7.915$, $p = 0.012$. The effect is plotted in Figure 10, which shows LC pre and post.

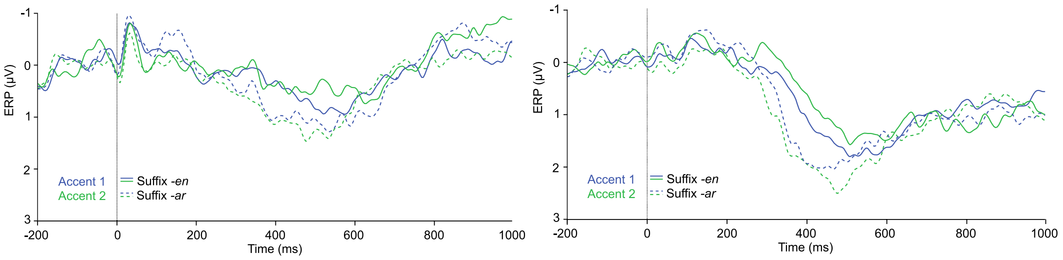


Figure 10. Left Central electrode site. 0 = suffix onset. The left one shows pre training and the right one shows post training. Blue line = Accent 1. Green line = Accent 2. Solid line = Suffix *-en*. Broken line = Suffix *-ar*.

No P600 effect was found in the time window 400-600 ms.

4.2 Behavioral data

Accuracy in the pre- and post-test was overall high. The results are shown in Table 16. There was no significant increase in accuracy for any condition.

Table 16. Mean accuracy in the pre- and post-test.

	Mean accuracy (*SD*) in %
Pre-test	93.78 (*17.60*)
Post-test	96.71 (*5.32*)

As for response times, the results are shown in Table 17.

Table 17. Response times in the pre- and post-test.

			Mean response time (*SD*) in ms
Pre-test	**Accent 1**	**Valid**	1026.22 (*336.27*)
		Invalid	1004.89 (*321.59*)
	Accent 2	**Valid**	1019.93 (*313.77*)
		Invalid	1035.30 (*332.65*)
Post-test	**Accent 1**	**Valid**	709.78 (*361.63*)
		Invalid	659.82 (*340.68*)
	Accent 1	**Valid**	668.79 (*337.47*)
		Invalid	709.88 (*346.05*)

Response times were significantly shorter for all conditions post training, $F(1,18) = 18.294$, $p < 0.001$. There was also an effect for suffix post training, where the response times for the plural suffix yielded significantly shorter response times (accent 1 invalid, accent 2 valid), $F(1,18) = 5.82$, $p = 0.03$.

4.3 Production data

There was moderate inter-rater agreement, with 72 % consistency, K = 0.42. The major difference between the two raters was found in the category neither accent/unsure vs. focused accent 1, where the author had annotated 15 % of the material as neither accent/unsure, and the controller had annotated 61 % of the material as neither/unsure. Instead, the author had annotated 56 % of the material as focused accent 1, and the controller 25 %. Interestingly, they were almost in complete agreement on focused accent 2 words, where the author had annotated 9 % of the material as accent 2 and the controller 10 %. The discrepancies can be explained by the notion that the author had a deeper knowledge of the material, and also was an L1 speaker of Central Swedish, whereas the controller was an L1 speaker of South Swedish, and might therefore have shown less certainty in his judgments. This motivated the use of the author's annotations despite the differences.

Accuracy significantly increased for both accents post game, as seen in a main effect for pre-post, $F(1,18) = 11.47$, $p = 0.003$. There was also a main effect for accent, where accent 1 had significantly higher accuracy than accent 2, $F(1,18) = 22.86$, $p < 0.001$. Descriptive statistics are shown in Table 18. There were no interactions, neither for accent, nor game presence.

Table 18. Mean accuracy in the production test per accent pre and post training.

		Mean accuracy (*SD*) in %
Pre training	**Accent 1**	68.75 (*34.79*)
	Accent 2	16.12 (*27.15*)
Post training	**Accent 1**	73.18 (*30.13*)
	Accent 2	26.00 (*30.36*)

As seen in the standard deviation the variation was large, and some participants did not show any increase at all, whereas some did. Figure 11 shows a confidence interval plot where some overlap between the sessions is visible.

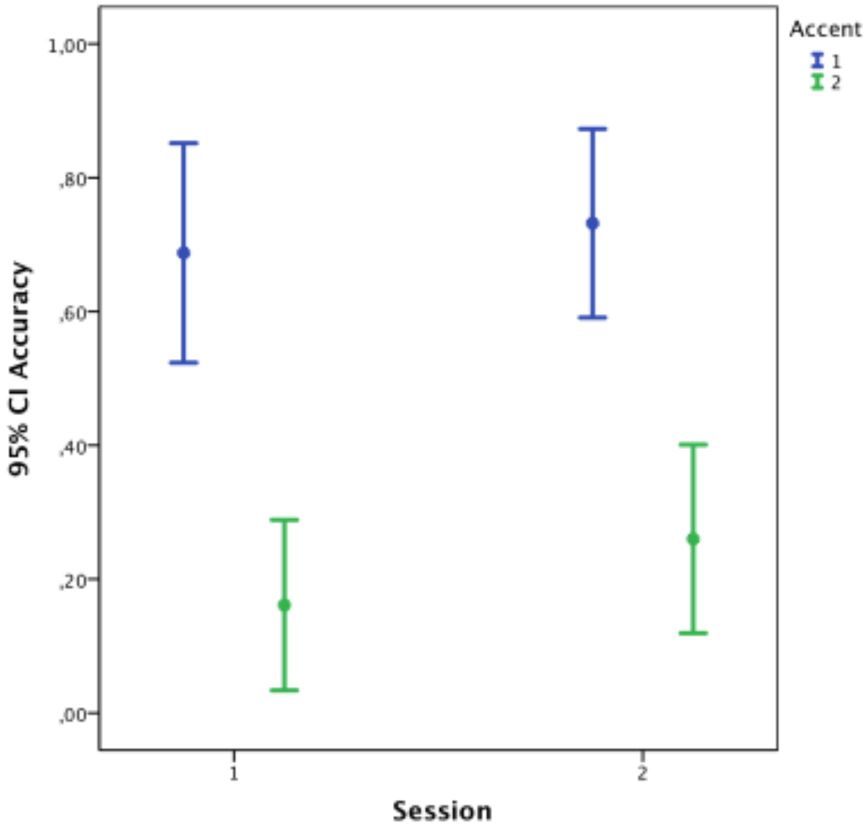


Figure 11. Confidence interval plot of production accuracy.

To see what the increase in accuracy can look like, before and after examples of F0-curves from one participant whose accent 2 accuracy went from 0 to 75 % are shown in Figure 12.

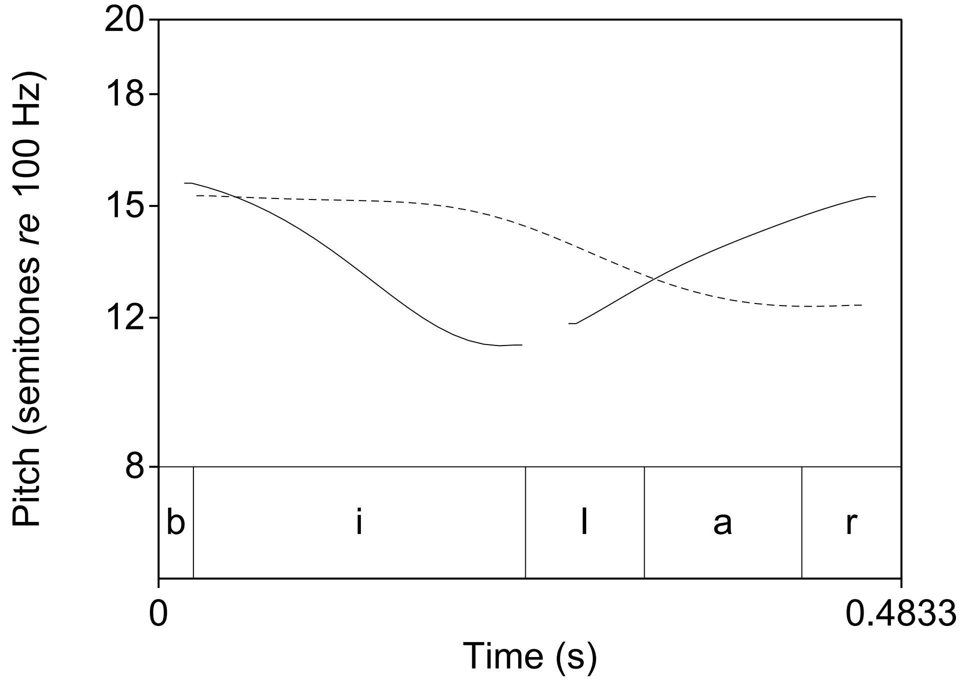


Figure 12. Accent 2 target word pre and post training. Dashed line = pre training. Solid line = post training.

Another aspect found in the data is what seems to be South Swedish like accent 2, characterized by a later peak in the accent. This was found sporadically in three different participants post training, and one example from a speaker with Italian L1 is given in Figure 13. However, the same speaker also had an increase in accuracy for Central Swedish accent 2.

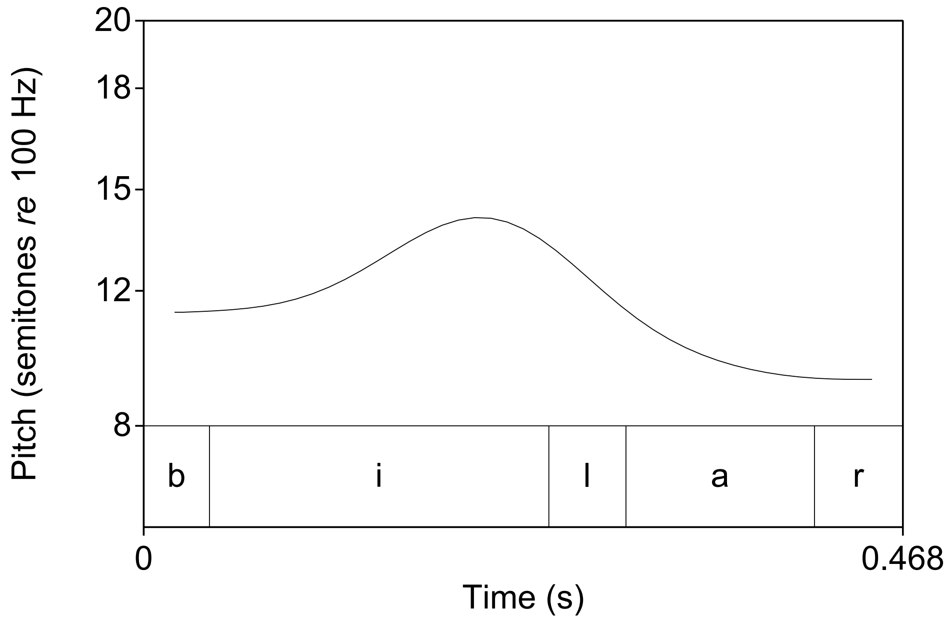


Figure 13. South Swedish like Accent 2 produced by one of the participants post training.

4.4 Game data

There was a negative correlation between stage and response time, r = .304, $p <$ 0.001, showing that response times decreased during the stretch of the game. Figure 14 shows response time averages per stage for correct responses for all participants.

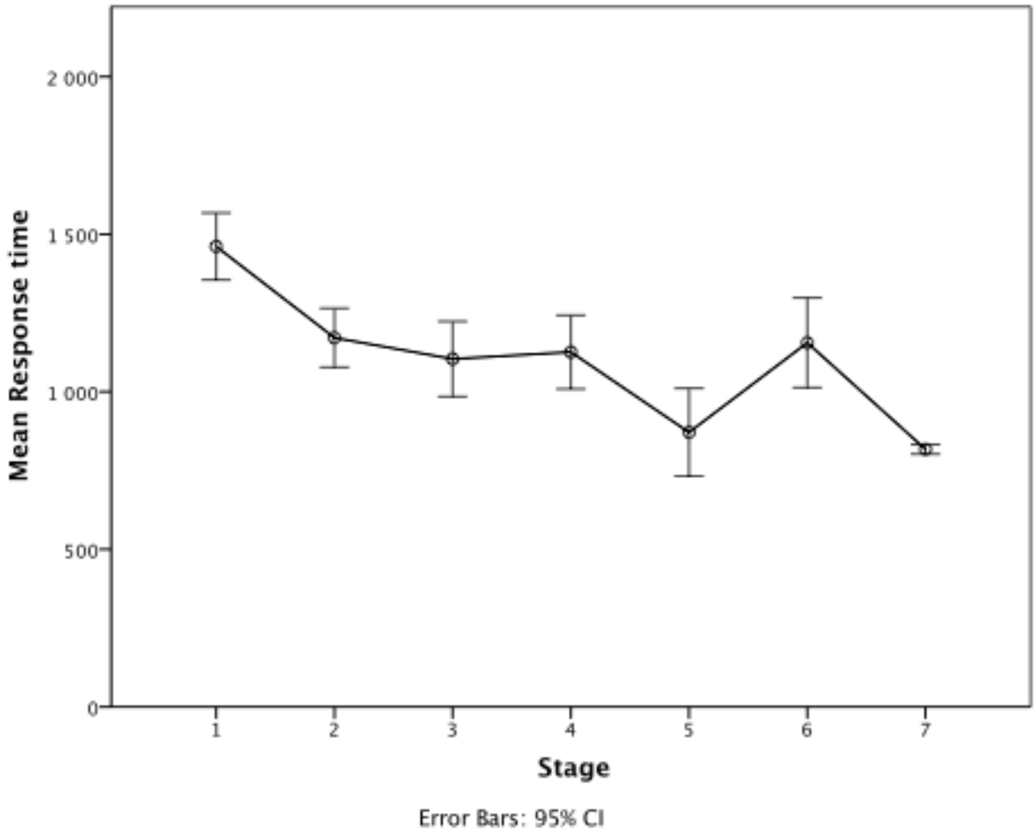


Figure 14. Mean response time for correct responses per stage in the Language Melody Game.

There was also a correlation between stage and accuracy, $r = 0.340$, $p < 0.001$, meaning that the accuracy significantly increased for each stage in the game. Figure 15 shows accuracy per stage for all participants.

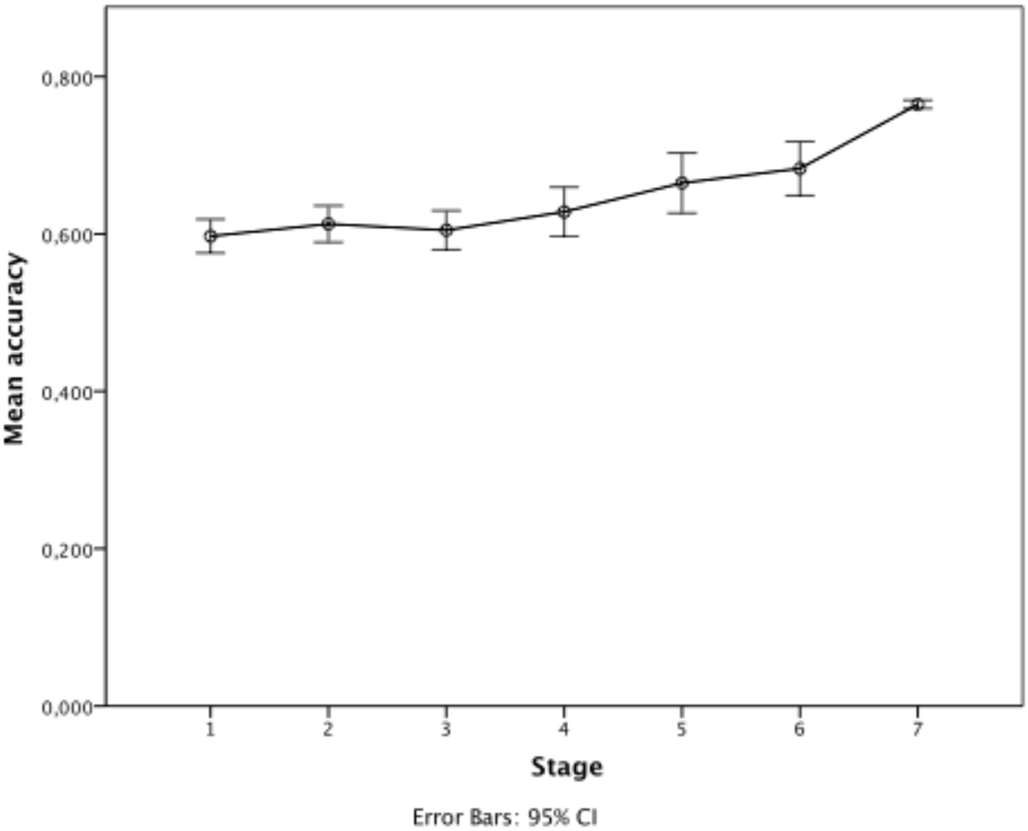


Figure 15. Mean accuracy per stage in The Language Melody Game.

The average amount of playtime was 7.48 h, (*SD = 2.58*). The mean accuracy on the last 5 rounds of level 7 for all participants was 77.74 % (*SD = 12.43*).

4.5 Correlations

The final accuracy correlated positively with the proposed LAN, $r = .555$, $p = 0.014$. This means that the more reduced (more positive) the LAN was, the higher the final accuracy was in the game. This can be seen plotted in Figure 16. Of the other tested correlations, none was significant.

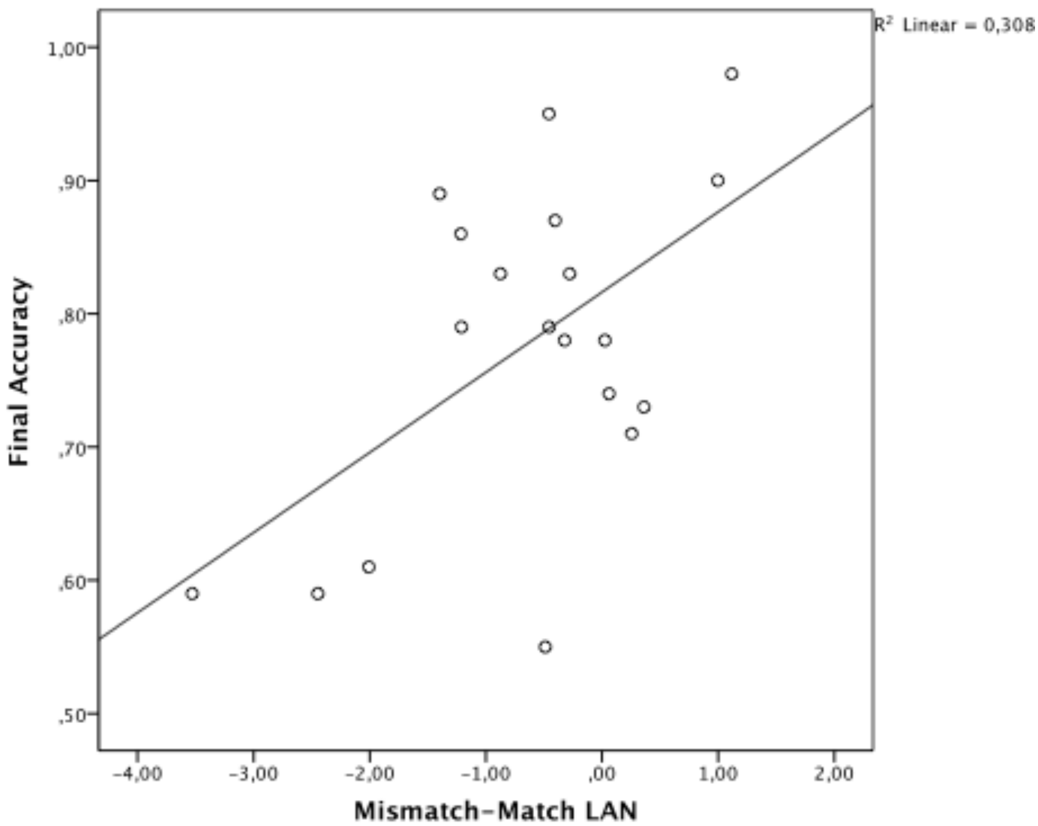


Figure 16. Correlation between final accuracy and Invalid-Valid LAN.

5 Discussion

In this thesis, a study on L2 acquisition of the Central Swedish word accent – suffix association has been presented. Low to intermediate L2 learners of Swedish trained this association in an online game for two weeks. In the game, they were asked which suffix should follow after hearing the word accent on the stem of a target word. A main idea to the game is that it promotes implicit learning. The task does not focus on *which word accent a particular word has*, but rather *what the word accent implies,* and *what suffix is associated with it*. This can be argued to be more ecologically valid than training identification or discrimination of word accents, since this comes more closely to how L1 speakers actually use word accents. Before and after playing the game, the participants took part in a production and perception test, and the perception test also included an EEG recording where ERPs were elicited. The main question posed in this study concerned how perception, production and neural activity would be affected by a two-week training period of the word accent – suffix association for L2 learners of Swedish. Follow up questions concerned whether perception, production and neural activity would be more L1-like after training, and whether perception training alone would increase accuracy in production.

Starting with the neural activity, two ERP onsets were investigated: word accent onset and suffix onset. In the word accent ERPs there was increased negativity anteriorly and centrally post game in the pre-activation negativity (PrAN) time window 136-280 ms after accent onset. This increased negativity indicates that the participants did make increased use of the word accents as predictors after the two-week period. This was also in line with the predictions for this study. There was also a larger negativity for accent 1 in the PrAN window, both pre and post training. This is the same tendency that has been found for Swedish L1 speakers (Söderström et al., submitted; Roll et al., 2015; Roll, 2015). However, there was no relatively greater increase in the accent 1 PrAN post game. Both accents yielded increased negativity.

The PrAN finding indicates that the participants did distinguish between the word accents. It shows that they had implicitly learned to use them as predictors to a certain extent, even before training the association. It also shows that they had learned that accent 1 has fewer alternatives for endings than accent 2. PrAN correlates with type frequency, increasing for accent 1, which has fewer types than accent 2. This has been explained by the notion that accent 2 has almost infinite possibilities of endings since it is the accent used in compounds in Central Swedish, and compound productivity is very high in Swedish. Therefore, in a follow up study, it would be interesting to include compounds in the game as well to see how this would affect the results. In such a study the difference between accent 1 and 2 would be expected to increase post training.

The findings of a PrAN effect pre training can also be compared to Gosselke Berthelsen et al. (in preparation), who did not find any PrAN effect for low to intermediate L2 learners of Swedish at all. The results from this study are thereby not in line with their study, and might be explained by a possible higher proficiency of the participants in this study compared to those in Gosselke Berthelsen et al. (in preparation).

Following Wang (2003b), there is a notion that tonal processing is right lateralized for L2 learners with non-tonal L1s, but gets more left lateralized with higher proficiency, as has been shown for L2 learners of Mandarin Chinese. No laterality effect was found for word accents in this study, however. Comparing the results of the L2 learners to results from Swedish L1 speakers, a left lateralized PrAN is found in them (Söderström et al., submitted; Roll, 2015). The location of the processing is, as stated in section 2.2, not possible to tell from EEG data, since it does not provide information about the source of the electrical activity, so we still do not know much about the laterality of the processing of Swedish word accents in L2 learners. Locality of the processing, along with plasticity and changes in brain structure could be studied using magnetic resonance imaging (MRI) in a follow-up study.

In the suffix onset ERPs both similarities and dissimilarities from L1 speakers were found. A left anterior negativity (LAN) effect was found for invalid accent/suffix stimuli post training in the Left Anterior ROI at 225-300 ms post suffix onset, but not pre training. This could indicate that the participants implicitly learned the grammatical rule of accent assignment. Comparing this result to L1 speakers, Söderström et al. (in press) found a LAN effect for L1 speakers in the same kind of task. One interesting point is that Söderström et al. used pseudowords in their study, and that is the only study on L1 processing of accent – suffix combinations that has reported a LAN. This might indicate that a LAN is found where morphology is more in focus than lexical content, since the pseudowords lack thereof. It is possible that a pseudoword task yields a different kind of processing which is more morphologically driven, where the words are processed as decomposed units of morphemes, compared to a study using existing words, which possibly could be stored as whole units, as well as decomposed morphologically. If the participants did not know the meaning of the words, there is a possibility that the words used in the experiment functioned as pseudowords for them. Unfortunately, no control was made on knowledge of the words used, and there is only anecdotal evidence for this, since some participants asked what the words meant during the experiment session. It was expected that LAN would correlate with higher final accuracy in the game, and that would have indicated that the participants who learned the association better would react more strongly to the invalid accent – suffix combinations. However, the opposite result was found. A reduction in LAN correlated with higher accuracy. If LAN indexes morphological processing, this might imply that participants with lesser knowledge of the words decomposed them more into separate morphological units and therefore showed an enhanced LAN. Participants with higher accuracy in the game might also have had a better knowledge of the Swedish lexicon and thereby had developed memory representations for the inflected words as whole units.

This might have led to these participants processing the words as whole units instead of separate morphological units, and hence, yielding a smaller LAN.

One thing that differs between the L2 learners and the previous findings in L1 learners is that the participants showed an effect for suffix and not validity in the post-test 325-425 ms after suffix onset, where the plural suffix yielded more positivity in the post-test. This indicates that the L2 learners were still more focused on segmental phonology, which is more salient, than on intonation. This could be explained by the fact that none of the participants had tone as a phonological feature in their L1, and thereby had no bias to react to the tones more. With even more training and proficiency in Swedish the results might resemble L1 processing more. An interesting follow up to this would be to compare the results from this study to a study on L1 speakers of tonal languages learning Swedish, participating in the same pre- and post-test and training. Since previous research has pointed towards the notion that L1 learners of tonal languages learn tones better in an L2 this could also be a possibility in this case. There is evidence both from tonal languages (Francis et al., 2008; Kaan et al., 2007) and Norwegian, a language closely related to Swedish, where Van Dommelen & Husby (2009) found that Mandarin L1 speakers learned the word accents in Norwegian better than their German counterparts. Hence, it would be interesting to conduct a study with these premises. It would also be interesting to look at L2 learners with word accents in their L1 since some tendencies have been presented, indicating that they learn the word accents easier in Swedish (Tronnier & Zetterholm, 2013) and make a comparison between the different typological tonal features in the L1 and investigate how this affects L2 learning. Another notion about the participants in this study is that the rhythmical structure of their L1s has not been taken under consideration, and this could have affected the results. In a further study, a division between participants with stress-timed languages (e.g. German, English) and more syllable-timed languages (e.g. French, Spanish) could be made.

No P600 was found for invalidity post training, as found in L1 speakers (Söderström et al., in press; Roll, 2015), which indicates that the participants did not reprocess the ungrammaticality of the word when hearing the suffix made invalid by the word accent. These results are opposite of what Rossi et al. (2006) found. In that study, a P600 was found for low to intermediate learners in an experiment concerning morphology, but no LAN. The LAN was visible for learners on higher level along with a P600. In this study, the opposite was found. There was a LAN effect, but no P600. However, this study was not on morphology alone, but on the integration of tonal phonology and morphology, and the discrepancy between the present and previous results might be due to that notion. One possible explanation for the presence of LAN but lack of P600 might be that the participants did notice the mismatch, but did not reprocess the ungrammaticality. The PrAN effect also corroborates the use of the word accents as predictors. As visible in the ERPs and the response times in the behavioral data there was an effect of suffix, indicating that the suffixes gave the participants more useful information than the accents. This might thereby imply that they did use the word accents as predictors, and noticed the mismatch, but they still relied more on the suffixes, and no reprocessing was executed. Further studies of word accent processing in L2 learners are needed to compare these results and draw firm conclusions.

Concerning the behavioral measures in the pre- and post-test, the participants' overall response times did decrease, indicating that they had learned the suffixes better. However, they did not show L1-like responses after the training period. Where L1 speakers have shorter response times for valid accent - suffix combinations, the L2 learners had shorter response times for the plural suffix, regardless of preceding accent. This is also in line with the findings from the ERP data, which showed an effect of suffix. An explanation for the shorter response times for the plural suffix is that it might be due to singular/plural ambiguity of the suffix *–en*. In the paradigm in this study *-en* functioned as a singular suffix, but in another noun paradigm in Swedish it functions as a plural suffix. The plural suffix in this study *–ar* is always a plural suffix, and therefore might have yielded shorter response times since there is no ambiguity and thereby no confusion about this. On an anecdotal notice, some of the participants asked about this matter during the experiment. The response times can be compared to the experiment conducted in Schremm et al. (2016), where fairly proficient L2 learners of Swedish showed the same tendency in response times as L1 speakers, with longer response times for invalid word accent - suffix combinations. The two weeks of special training for the low to intermediate L2 learners in this study did not show to be sufficient to acquire the same kind of results as in Schremm et al.'s experiment. For further research, it would be interesting to see a study with a longer training period, to see how this would affect the results, and if the participants then would become more L1-like in their results.

Comparing ERPs and behavioral data, where the ERP data showed signs of prediction, but the behavioral data did not, this can be discussed in the light of van Hell and Tokowicz (2010). They argue that learning can be seen earlier in the neural response than when it is seen in behavioral data, which is exactly what is found in this study. PrAN and LAN effects indicated that the participants had learned to use word accents predictively, but in the pre- and post-test the behavioral results indicated that they had not. However, for the game data, response times decreased, and accuracy increased during the stretch of the game. This is in line with what was expected and indicates that the participants did get better at the accent - suffix combinations. These results, however, have not transferred into the task in the pre- and post-test.

When it comes to production, the participants showed an increase in accuracy of the word accents, which is in line with Wang et al. (2003a) who showed the same tendency for English L1 speakers training perception of Mandarin tones. In Wang et al.'s study, the participants also increased their accuracy in producing the tones even if there was no explicit production training. The increase in production shows that this can be learned implicitly, without special training in production. However, the variation in this study was large, as shown in Figure 11. No possible explanation for this variation has been found, since there were no correlations between production accuracy and playtime, final accuracy in the game or LAN. Some participants' results vastly improved, whereas others' results did not improve at all. There was no effect of game presence in the results, meaning that the production results improved for words that were present in both the production test and the game, as well as for words that were not present in the game but were new to the production test. The rule of accent/suffix assignment was thereby generalized onto novel words, and this finding suggests that the participants learned the rule of accent assignment, rather than learning the intonation of each word as a whole unit.

As for the inter-rater agreement, it was only moderate. It can however, be explained by the fact that the author had better knowledge of the material, and that the controller did not have Central Swedish word accents in his variety of Swedish. The decision to still go with the author's annotations in the analysis was motivated by this notion. Interestingly, the two raters were in almost complete agreement about the focused accent 2, which is probably the most salient intonational contour with its two peaks.

One interesting phenomenon found in the material is that sporadically, three participants produced what resembled South Swedish word accents, although the game only consisted of Central Swedish accents. The participants are surrounded by South Swedish, which is the variety traditionally spoken in Lund where their studies take place. Hence, acquisition of South Swedish word accents might be explained by increased awareness of word accents as a phonological phenomenon and incorporation of the variety of Swedish spoken in their surroundings. One critique of the study is that the participants learned Central Swedish word accents although they lived in a South Swedish speaking area. There might have been more ecological validity if they had learned the South Swedish accent patterns. However, Lund is a city with a large number of students from all over Sweden and Central Swedish is probably also a part of the participants' linguistic surroundings. Central Swedish is also what is mostly heard on the radio and television. The findings in this study, however, show that it is possible to learn a variety of Swedish other than the one spoken in the area.

To summarize the discussion, what has been shown is that a PrAN effect with increased overall negativity in the post-test, a LAN effect, the game results and the production results indicate that the L2 learners did increase their use of word accent – suffix association after playing the game for two weeks. They became L1-like concerning the PrAN and the LAN but not in the lack of a P600 effect, and the production accuracy did not reach L1 level. Furthermore, the suffix effect in the ERPs along with the behavioral measures from the pre- and post-test showed that they had not acquired the association enough to achieve the same results as more proficient L2 learners and L1 speakers of Swedish. However, results show a reliable tendency towards more L1-like patterns, which could be expected to continue if longer training were administered.

6 Conclusions

A little more is now known about the processes involved in L2 acquisition of the Swedish word accent - suffix association. After two weeks of word accent training, L2 learners of Swedish showed significant changes in perception, production and neural activity. ERP data showed, with an increased PrAN effect and the emergence of a LAN effect, that the participants had started to use word accents as predictors. The game data also showed, with higher accuracy and shorter response times, that the participants did acquire the association rule. Furthermore, the results of the production test showed that the participants increasingly used word accents in their speech as well.

The post training results were in some ways more L1-like than pre training. The increased PrAN effect and the greater negativity for accent 1 have also been found in L1 speakers, along with a LAN. However, LAN for invalidity has only been reported in a study using pseudowords. An explanation for this might be that in a pseudoword experiment, no lexical content is given and the participant has to rely only on rules associating tone with morphological information. In an experiment with real words, the participant might be able to use *both* the tone-morphology-associating rules *and* the whole word as a stored unit. An interpretation of this is that a LAN is found where the participant needs to decompose the word and process the morphology, and not if the word is stored as a whole unit. The LAN found in this study could indicate that the participants might have had low knowledge of the meaning of the words in the study. If they did not know the meaning of the words, the words would function as pseudowords. This might also explain the opposite correlation of LAN with final accuracy. Participants who performed with higher accuracy in the game might have had better knowledge of Swedish, and hence, better lexical knowledge. This would then suggest that the participants with better lexical knowledge were more prone to process words as whole units, and thereby, showed a smaller LAN. To go to the bottom of this, both for L1 speakers and L2 speakers, further research is needed.

The lack of a P600 along with the response times from the pre- and post-test show discrepancies between the L1 and L2 speakers. In experiments with L1 speakers, invalid conditions yield a P600, which means that they reprocess the ungrammaticality of the invalid condition, and this is shown in response times as well, where invalid conditions yield longer response times. This was not found in L2 speakers. Instead, there was no P600 and a suffix effect in the post-test, where the plural suffix yielded significantly shorter response times. These results seem to imply that the participants pre-activated suffixes, with accent 1 as a stronger predictor; they reacted to the mismatching condition, but still relied more on the suffix information and thereby did not reprocess the ungrammaticality.

As for the production results, they showed that the perception-only training was indeed effective for enhancing production, where the post-test results showed an increased correct use of word accents. The rule of accent assignment was also generalized onto novel words. However, the variation was considerable, and the cause of this has not been explained by data provided from this study.

For further research, it would be interesting to dig deeper into L2 acquisition of word accents. More L1 backgrounds could be taken into account (word accent languages, tone languages, different rhythmical structures) along with longer training periods. Other methods, such as MRI could give more insight into what happens in the brain during acquisition. With MRI, locality of processing and plasticity can be investigated, leading to better knowledge of L2 processing as well as the differences and similarities between L1 and L2 processing of the word accent – suffix association. It would also be interesting to see studies on L1 acquisition of word accents, since, to my knowledge, there is no neurolinguistic studies focusing on this. Another interesting point to look into, as mentioned above, is the function of the LAN effect, and the impact and interplay of intonation, morphology and lexical information.

7 References

Boersma, P. & Weenink, D. 2014. *Praat – doing phonetics by computer. Version: 5.4.08*. Computer program. Retrieved at: http://www.praat.org November 2014.

Bruce, G. 1977. *Swedish Word Accents in Sentence Perspective*. Lund: Gleerup.

Bruce, G. 1987. How floating is focal accent? In: Gregersen, K., Basbøll, H. (eds.) *Nordic Prosody IV*, Odense University Press, Odense, 41-49.

Bruce, G. & Gårding, E. 1978. A Prosodic Typology for Swedish Dialects. In: Gårding, E., Bruce, G., Bannert, R. (eds.), *Nordic Prosody*. Lund: Gleerup, 219–228.

Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement 20*, 37-46.

Ejerhed E., Källgren G., Wennstedt O. & Åström M. 1992. The linguistic annotation system of the Stockholm-Umeå Corpus project. *Technical Report 33*, Department of General Linguistics, Umeå University, Umeå.

Francis, A.L., Ciocca, V., Ma, L. & Fenn, K. 2008. Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics 36*, 268-294.

Gosselke Berthelsen, S., Horne, M., Shtyrov, Y., Brännström, J. & Roll, M. In preparation. Early second language learners' processing of pre-activating prosodic cues: An electrophysiological study of Swedish word accents.

Hed, A. 2014. Production and Perception of Swedish Word Accents by Somali L1 speakers. BA Thesis. Lund University.

Kaiser, R. 2011. Do Germans produce and perceive the Swedish word accent contrast? A cross-language analysis. *TMH-QPSR, 51(1),* 93-96.

Kaan, E., Wayland, R., Bao M. & Barkley, C. M. 2007. Effects of native language and training on lexical tone perception: An event-related potential study. *Brain Research 1148*, 113-122.

Klein, D., Zatorre, R.J., Lee, K.M. & Hirsch, J. 2001. A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *NeuroImage 13*, 646-653.

Lahiri, A., Wetterlin, A. & Jönsson-Steiner, E. 2005. Lexical Specification of Tone in North Germanic. *Nordic Journal of Linguistics 28*, 61-96.

Lu, S., Wayland, R. & Kaan, E. 2015. Effects of production training and perception training on lexical tone perception – A behavioral and ERP study. *Brain Research 1624*, 28-44.

Luck, S.J. 2005. *An Introduction to the Event-Related Potential Technique*. Cambridge: The MIT Press.

Myrberg, S. 2010. *The Intonational Phonology of Stockholm Swedish*. Stockholm Studies in Scandinavian Philology.

Myrberg, S. & Riad, T. 2015. The prosodic hierarchy of Swedish. *Nordic Journal of Linguistics, 38*,115-147.

Riad, T. 2009. Prosodi i svenskans ordbildning och ordböjning. Ms., Stockholm University.

Roll, M. 2006. Prosodic cues to the syntactic structure of subordinate clauses in Swedish. In: Bruce, G. & Horne, M. (eds.), *Nordic Prosody: Proceedings of the IXth Conference, Lund 2004*, 195-204.

Roll, M. 2015. A neurolinguistic study of South Swedish word accents: Electrical brain potentials in nouns and verbs. In: *Nordic Journal of Linguistics, vol 38* (2)*,* 149-162.

Roll, M., Horne M. & Lindgren M. 2010. Word accents and morphology – ERPs of Swedish word processing. *Brain Research 1330*, 114-123.

Roll, M., Horne, M. & Söderström P. 2013. Word-stem tones cue suffixes in the brain. *Brain Research 1520*, 116–120.

Roll, M., Söderström, P., Mannfolk, P., Shtyrov, Y., Johansson, M., van Westen, D. & Horne, M. 2015. Word tones cueing morphosyntactic structure: Neuroanatomical substrates and activation time-course assessed by EEG and fMRI. *Brain and Language, 150*, 14-21.

Rossi, S., Gugler M.F., Friederici, A.D. & Hahne, A. 2006. The impact of proficiency on syntactic second-language processing of German and Italian: evidence from event-related potentials. *Journal of Cognitive Neuroscience 18*, 173-204.

Schremm, A., Söderström, P., Horne, M. & Roll, M. 2016. Implicit acquisition of tone-suffix connections in L2 learners of Swedish. *The Mental Lexicon 11:1,* 55–75.

Söderstrom, P., Roll, M. & Horne, M. 2012. Processing morphologically conditioned word accents. *The Mental Lexicon 7:1*, 77–89.

Söderström, P., Horne, M., Frid, J. & Roll, M. (submitted). Pre-activation negativity (PrAN) in brain potentials to unfolding words.

Söderström, P., Horne, M. & Roll, M. (in press). Stem tones pre-activate suffixes in the brain. *Journal of Psycholinguistic Research*.

Tronnier, M. & Zetterholm E. 2013. Tendencies of Swedish word accent production by L2 Learners with tonal and non-tonal L1. In Eva Liina Asu, Partel Lippus (eds.) *Nordic Prosody, Proceedings of the XIth Conference, Tartu 2012*, 391-400.

Tronnier, M. & Zetterholm, E. 2014. Appropriate tone accent production in L2 Swedish by L1 speakers of Somali? C*oncordia Working Papers in Applied Linguistics: Proceedings of New Sounds*, *Montreal*, 722-736.

Van Dommelen, W. & Husby, O. 2009. Perception of Norwegian word tones by Chinese and German listeners. In M. A. Watkins, A. S. Rauber & B. O. Babtista (eds.), *Recent research in second language phonetics/phonology: Perception and production*, 308-321. Newcastle upon Tyne: Cambridge Scholars Publishing.

van Hell, J.G. & Tokowicz, N. 2010. Event-related brain potentials and second language learning: syntactic processing in late L2 learners at different L2 proficiency levels. *Second Language Research 26,1,* 43-74.

Wang, Y., Jongman, A. & Sereno J. 2003a. Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustic Society of America* 113 (2), 1033-1044.

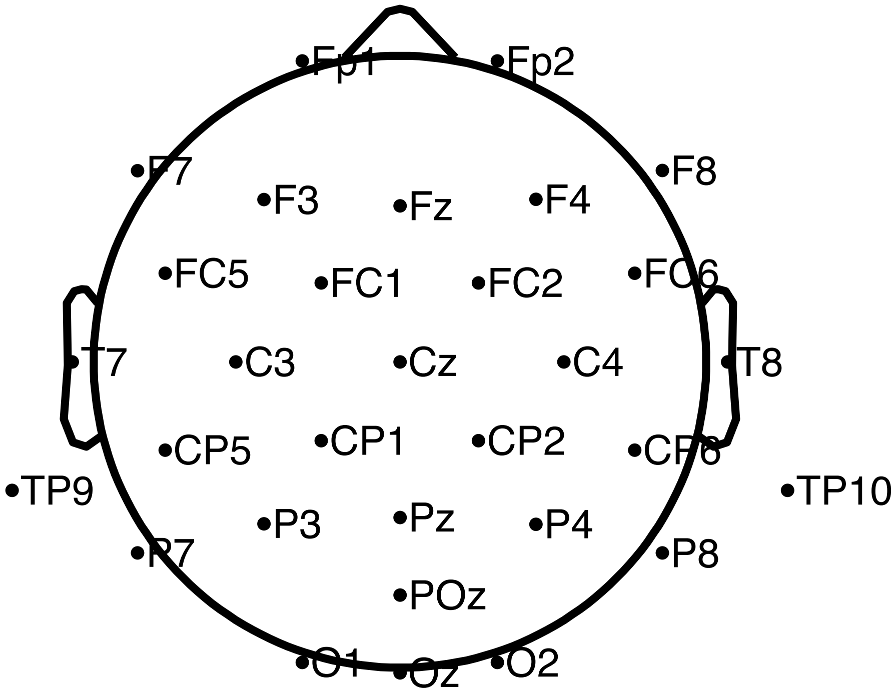
Wang, Y., Spence, M.M., Jongman, A. & Sereno, J. 1999. Training American listeners to perceive Mandarin tones. In: *Journal of the Acoustic Society of America* 106 (6), 3649-3658.

Wang, Y., Jongman, A. & Sereno, J. & Hirsch, J. 2003b. fMRI evidence for cortical modification during learning of Mandarin lexical tone. *Journal of Cognitive Neuroscience, 15:7*, 1-9.

Wayland, R.P. & Guion, S.G. 2004. Training English and Chinese Listeners to Perceive Thai Tones: A Preliminary Report. *Language Learning 54:4*, 681-712.

Weyerts, H., Penke, M., Dohrn, U., Clahsen, H. & Münte. T.F. 1997. Brain potentials indicate differences between regular and irregular German plurals. *NeuroReport 8*, 957-962.

Appendix 1. Electrode setup



Appendix 2. LMG Stimuli words

NOUNS	*Type*	*Token*
	bil ’car’	5
	bild ’picture’	5
	chef ’boss’	5
	dag ’day’	5
	dörr ’door’	5
	film ’movie’	5
	form ’form’	5
	färg ’color’	5
	hund ’dog’	5
	häst ’horse’	5
	kropp ’body’	5
	kväll ’evening’	5
	plats ’place’	5
	skog ’forest’	5
	säng ’bed’	5
	text ’text’	5
	väg ’road’	5
	vägg ’wall’	5
	vän ’friend’	5
	värld ’world’	5
Sum:	**20**	**100**

VERBS	*Type*	*Token*
	bygga ’to build’	6
	läsa ’to read’	6
	resa ’to travel’	6
	ringa ’to call’	6
	drömma ’to dream’	5
	gräva ’to dig’	5
	leka ’to play’	5
	rymma ’to escape’	5
	glömma ’to forget’	4
	tänka ’to think’	4
	leva ’to live’	3
	röka ’to smoke’	3

	bränna ’to burn’	2
	dröja ’to linger’	2
	döma ’to judge’	2
	känna ’to feel’	2
	skölja ’to rinse’	2
	släppa ’to let go’	2
	vända ’to turn around’	2
	värma ’to heat’	2
	växa ’to grow’	2
	böja ’to bend’	1
	gifta ’to marry’	1
	gömma ’to hide’	1
	hjälpa ’to help’	1
	klämma ’to squeeze’	1
	krympa ’to shrink’	1
	kräva ’to demand’	1
	kyssa ’to kiss’	1
	lyfta ’to lift’	1
	märka ’to notice’	1
	nämna ’to mention’	1
	nöja ’to be content’	1
	skilja ’to divorce’	1
	skrämma ’to frighten’	1
	skälla ’to scold’	1
	skärpa ’to shape up’	1
	slänga ’to throw’	1
	spränga ’to blow up’	1
	sträcka ’to stretch’	1
	stämma ’to sue’	1
	svänga ’to turn’	1
	söka ’to seek’	1
	tycka ’to think’	1
	väcka ’to wake up’	1
Summa:	**45**	**100**

Appendix 3. Production sentence list 1.

s	4	Var såg Anna dagar?	Anna såg dagar på TV.
v	3	Vad glömmer Tom?	Tom glömmer saker.
s	9	Var såg Anna granen?	Anna såg granen på TV.
s	1	Var såg Anna bilen?	Anna såg bilen på TV.
s	6	Var såg Anna muggar?	Anna såg muggar på TV.
v	9	Vad väger Tom?	Tom väger saker.
v	2	Vad byggde Tom?	Tom byggde saker.
s	3	Var såg Anna chefen?	Anna såg chefen på TV.
v	1	Vad bränner Tom?	Tom bränner saker.
v	5	Vad sköljer Tom?	Tom sköljer saker.
s	7	Var såg Anna salen?	Anna såg salen på TV.
v	7	Vad döper Tom?	Tom döper saker.
s	8	Var såg Anna skedar?	Anna såg skedar på TV.
v	10	Vad ägde Tom?	Tom ägde saker.
s	2	Var såg Anna bilder?	Anna såg bilder på TV.
s	5	Var såg Anna dörren?	Anna såg dörren på TV.
s	10	Var såg Anna knivar?	Anna såg knivar på TV.
v	8	Vad klippte Tom?	Tom klippte saker.
v	6	Vad sände Tom?	Tom sände saker.
v	4	Vad krympte Tom?	Tom krympte saker.

Appendix 4. Production sentence list 2.

v	8	Vad klipper Tom?	Tom klipper saker.
s	2	Var såg Anna bilden?	Anna såg bilden på TV.
s	5	Var såg Anna dörrar?	Anna såg dörrar på TV.
v	3	Vad glömde Tom?	Tom glömde saker.
v	1	Vad brände Tom?	Tom brände saker.
v	5	Vad sköljde Tom?	Tom sköljde saker.
s	7	Var såg Anna salar?	Anna såg salar på TV.
v	9	Vad vägde Tom?	Tom vägde saker.
s	1	Var såg Anna bilar?	Anna såg bilar på TV.
v	10	Vad äger Tom?	Tom äger saker.
s	3	Var såg Anna chefer?	Anna såg chefer på TV.
s	10	Var såg Anna kniven?	Anna såg kniven på TV.
s	9	Var såg Anna granar?	Anna såg granar på TV.
s	4	Var såg Anna dagen?	Anna såg dagen på TV.
v	2	Vad bygger Tom?	Tom bygger saker.
s	6	Var såg Anna muggen?	Anna såg muggen på TV.
s	8	Var såg Anna skeden?	Anna såg skeden på TV.
v	6	Vad sänder Tom?	Tom sänder saker.
v	7	Vad döpte Tom?	Tom döpte saker.
v	4	Vad krymper Tom?	Tom krymper saker.

Appendix 5. List of participants.

Participant	Age	Gender	L1	Level of Swedish	Time of Swedish in months	Time in Sweden in months
1	25	F	Finnish	B1	72	4
2	23	F	German	B1	24	3.5
3	22	F	Finnish	B1	84	3
4	22	M	Spanish	-	0	4
5	24	F	Dutch	A2	12	18
6	24	M	German	B1	7	3.5
7	23	F	Spanish	A2	2	3
8	21	F	Italian	B1	24	4
9	25	M	English	A2	5	4
10	23	F	Russian	B2	48	6
11	21	F	Russian	B2	48	6
12	20	F	French	-	0.5	6
13	26	M	German	A2	12	18
14	21	F	Spanish	A2	6	6
15	24	M	German	A2	18	0.5
16	21	F	German	A2	10	6
17	24	F	French	A2	36	1
18	21	M	English	A2	8	8
19	26	F	Russian/ Ukrainian	A2	5	6
Average (SD):	*23 (1.8)*				*22.2 (24.6)*	*5.8 (4.7)*