Neural Asymmetry in the Perception of South Swedish Word Accents

Evidence from Mismatch Negativity

Huayuan Cui (崔华元)

Supervisor: Prof. Mikael Roll

Centre for Language and Literature, Lund University
MA in Language and Linguistics, Phonetics
SPVR01 Language and Linguistics: Degree Project – Master's (Two Years) Thesis, 30 credits

May 2023

Abstract

South Swedish, as a dialect of modern Swedish, has two tonal word accents, *accent 1* and *accent 2*. Regarding these two word accents, there are three groups of opinion on which on is more lexically specified. The first group believes that accent 2 is more specified and accent 1 is default shaped by intonation, whereas the second group deems that accent 1 is lexically specified and accent 2 is the default accent. The third group, however, holds the opinion that both word accents are specified. In order to find evidence from brain level to support one of those opinions, a mismatch negativity (MMN) study under the passive oddball paradigm was conducted in the present study. Results show that in South Swedish accent 1 elicited significant early MMN while accent 2 elicited significant and robust late MMN. The asymmetry in temporal and amplitude domain suggests that accent 2 has more linguistics information encoded, which suggests that accent 2 in South Swedish has a more specific memory trace in native speakers' mental phonology. According to the underspecification theory, the more specified structure has more specific memory representation than underspecified features. In conclusion, the results of the present study support the first group's opinion, which is that accent 2 is more lexically specified than accent 1 in South Swedish.

Keywords: speech perception, South Swedish, Swedish word accents, mismatch negativity, MMN

Acknowledgements

First of all, I would like to thank my supervisor Prof. Mikael Roll for his instruction in many ways, from narrowing down the topic, programming for experiment and many valuable suggestions. Many thanks to Anna Hed and Renata Kochančikaitė for sharing me the Python code of a previous MMN experiment. Thank to Dr. Johan Frid as he helped me a lot on programming, without him I cannot even make the PsychoPy programme run. Thank Humanities Lab for providing the equipment for the experiment. Other credits go to Dr. Mechtild Tronnier and Prof. Tomas Riad from Stockholm University, they all contributed with suggestions to this study. I would also like to thank all my participants as they were all very kind and no one even complain a single word even if I messed with their hair using the gel. A special thank you to my speaker Joel Takahashi Olsson, without his perfect performance in the recording and his adequate Malmö accent it is impossible to complete this work. I am also willing to express my appreciation to Ann Hermansson, Sonja Holmer, Jinhee Kwon (권진희), Felix Lecaros and Nina Vang for helping me recruiting participants. Last but not the least, a big thank you to my biggest supporter in my life, Yanqing Huang (黄燕清), who is the one that gave me birth.

Table of contents

List of figures vi
List of tables vi
Abbreviations vii

Chapter 1 Introduction 1
1.1 Research questions and hypothesis 3
1.2 A snapshot of the method 3
1.3. Structure of the thesis 3

Chapter 2 Theoretical background 4
2.1 Swedish word accents 4
2.1.1 A brief introduction of Swedish word accents 4
2.1.2 A typological view of Swedish word accents 6
2.1.3 South Swedish word accents 6
2.1.4 Perception cues for Swedish word accents 7
2.1.5 Swedish word accents and its prediction function 8
2.2 EEG and ERP 8
2.3 MMN 9
2.3.1 Introduction of MMN 9
2.3.2 MMN and language processing 10
2.3.3 MMN asymmetry and its relation with phonology 10
2.4 Previous ERP studies on tonal patterns and Swedish word accents 11
2.4.1 Previous MMN studies 11
2.4.2 Other ERP studies on Swedish word accents 11
2.5 Tone processing in the brain 12

Chapter 3 Method 13
3.1 Experiment design 13
3.2 Stimuli recording and selection 14
3.3 EEG recording 15
3.3.1 Participants 15
3.3.2 Recording procedure 15
3.4 Data processing 16
3.5 Statistical analysis 17
3.6 Source localization 17

Chapter 4 Results and statistics 18
4.1 Results plotting 18
4.2 Results of statistical analysis 19
4.3 Results of source localization 22

Chapter 5 Discussion 24
5.1 Summarizing the asymmetry observed from the MMN results................................. 24
5.2 Identifying the late component in block 2 26
5.3 Evaluation of the method 27

Chapter 6 Conclusion..28

References..29

Appendix 34
Participant information 34

List of figures

Figure 1 Pitch contour of word accents in Central Swedish with accent 1 word ^{1}*manen* (the mane) in black and accent 2 word ^{2}*manar* (manes) in grey, sourced from Roll (2022, p. 02). Solid lines represent pitch contours of the focus accented level whereas dashed lines reflect the word accented level. The two prominence levels are introduced later in this section. 4

Figure 2 Geography distribution of Swedish word accents, sourced from Gårding and Lindblad (1973, p.48). 6

Figure 3 Pitch curve of two stimuli words, with ^{1}*kullen* on the left and ^{2}*kullen* on the right. 14

Figure 4 Channel locations (without ground channel AFz and ocular electrodes). 16

Figure 5 Average ERP of 19 participants at FC3 electrode with block 1 on the left and block 2 on the right. 18

Figure 6 Identity MMN of 19 participants at FC3 electrode with accent 1 on the left and accent 2 on the right. 19

Figure 7 Average deviant vs. average standard at FC3 electrode. 19

Figure 8 Topographies of significant electrodes distribution, with first peak at 254ms on the left and second peak at 350ms on the right. 20

Figure 9 Topography of significant electrodes distribution, with first peak at 168ms on the left and second peak at 592ms on the right. 21

Figure 10 Topography of significant electrodes distribution at 331ms on the left and 362ms on the right from accent 1 identity MMN. 21

Figure 11 Topography of significant electrodes distribution of accent MMN at 345ms............. 22

Figure 12 Source estimation of accent MMN at 345ms. 22

Figure 13 Source estimation of accent 1 vs. accent 2 MMN at 168ms. 23

Figure 14 Source estimation of accent 2 vs. accent 1 MMN at 350ms. 23

Figure 15 Accent 1 standard vs. accent 2 standard. 26

List of tables

Table 1 Prominence level for word accents, sourced from Riad (2014, p.184). 5

Table 2 South Swedish word accents. 7

Table 3 Experiment design. 14

Table 4 MMN comparison when accent 1 as deviant. 24

Table 5 MMN comparison when accent 2 as deviant. 25

Table 6 Peak and amplitude comparison. 26

Abbreviations

BA	Brodmann area
EEG	Electroencephalography
ERP	Event-related potential
F0	Fundamental frequency
H	High tone
Hz	Hertz
ICA	Independent component analysis
kΩ	kiloohm
L	Low tone
L1	First language
MMN	Mismatch negativity
ms	milliseconds
SD	Standard deviation
SOA	Stimulus onset asynchrony
μV	microvolt

Chapter 1 Introduction

Understanding speech is a challenging task to accomplish, requiring multiple procedures and cooperation between different parts of the visual-audio system. The variation of speech sounds and prosody require listeners to map the acoustic stimuli to sound representations in their mind. However, native speakers never seem to have difficulties during this procedure. Prosody plays an important role when talking about speech perception, for example, almost all languages use intonation to convey mood, emotion or mark focus information. Besides intonation, some languages use tone to determine meaning, like Mandarin Chinese. Some Scandinavian languages, including Swedish and Norwegian, use pitch accent to contrast meaning or facilitate word processing. Based on that, Swedish and Norwegian are sometimes regarded as languages with simple tonal systems. The main object language of the present study, South Swedish, which is a dialect of modern Swedish, is spoken in several provinces in southern Sweden. Similar to most other Swedish dialects, South Swedish has two tonal word accents, accent 1 and accent 2. Tonal accents in Swedish are sometimes used to contrast meaning (Elert 1972), facilitate lexical retrieval and use to predict word endings (Roll et al. 2010, 2015, 2017; Roll 2015, 2022; Söderström et al. 2016, 2017), which suggests that they have lexical functions to some extents. Hence, similar to phonemic features like consonants and vowels, some tones may have more of a specific memory trace since they are more specified according to underspecification theory (Archangeli 1988).

The central idea of underspecification theory in phonology is that only distinctive features are specified, which are represented in adults' long-term memory of sound, i.e., their phonology (Kiparsky 1985; Archangeli 1988). According to this theory, redundant features are not specified in the speaker's mental representation, which means they have less specific memory comparing to fully specified features. The asymmetry in long-term memory results in asymmetric perception. Over the years, several studies that focused on phonemic features have provided numerous empirical bases to support these theories. In the meantime, suprasegmental structures have often been neglected. However, some suprasegmental structures like tones in tonal languages have similar functions as phonemic features to certain extents. Thus, it is predictable that some tones have more specific memory representation than others because they are more specified. This suggestion has been supported by studies in some tonal languages (Kann et al. 2008; Politzer-Ahles et al. 2016). Politzer-Ahles et al. (2016) did an electroencephalography (EEG) study on Mandarin tone perception and concluded that in Mandarin Chinese there are asymmetries in the perception of tones. Based on their results, they argued that Mandarin tone 3 has less specific memory than other lexical tones because it only contains a subset of the features present in the memory representations of other tones due to its numeral allotonic and phonation type variation.

In the case of South Swedish word accents, the situation is a little bit complicated since there is no certain answer to which word accent is more specified. Over the years, Swedish phonologists constantly debate whether accent 1 or accent 2 is more marked or lexically specified. There are three groups of opinions, which I call Group 1, 2 and 3 in the following text. Group 1 holds the opinion that accent 1 is a default accent naturally assigned by intonation and accent 2 is more marked and lexically represented (Elert 1972; Riad 1998). Group 2 has a reverse opinion, in which disyllabic accent 1 is more marked and monosyllabic accent 1 and all accent 2 are assigned by default – a word that is lexically specified is assigned with accent 1 whereas for unspecified words with trochaic tree receives accent 2 and other unspecified words receive accent 1 (Lahiri, Wetterlin and Jönsson-Steiner 2005). Group 3 on the other hand, believe that both accents are lexically specified (Bruce 1977). Based on Group 1's opinion, accent 2 is more lexically specified, which indicates that accent 2 should have more specific memory in native speaker's mental lexicon. Based on group 2's opinion, accent 1 should have more specific memory trace than accent 2 because accent 1 is more lexically specified. Group 3's opinion, otherwise, predicted that both word accents have equal memory trace in the brain. Roll et al. (2010) described asymmetric perception in Central Swedish by conducting an EEG experiment, and argued that high Accent 2 stem tones are assigned morphophonologically by tone-inducing suffixes, supporting group 1. Felder et al. (2009) also conducted one study on asymmetric perception of Swedish word accents and their results were argued to support group 2's opinion. They conducted two behaviour tasks and assumed that accent 1 is more lexical specified based on reaction time and accuracy. However, the problem is whether reaction time and accuracy reflect lexical specification in mental lexicon as there are more factors may influence on processing time and accuracy such as more predictable word ending and word frequency etc. Roll (2015) criticized Felder et al. (2009) based on the fact that most accent 1 word they chose as stimuli only have one possible continuation after the first syllable, while accent 2 words have various possible continuations. Hence, in the present study, I am trying to find asymmetry in the perception of two word accents at the brain level using electroencephalography (EEG), trying to find evidence to support one of the three groups' hypothesis.

There are three reasons why South Swedish is the object of the present study. The first one is because the experiment is conducted in South Sweden, which is the homeland of South Swedish language. Second, compared to Central Swedish, there are fewer researches on South Swedish word accents, despite two word accents in South Swedish have distinct phonetic shape compare to Central Swedish. Last but not the least, there is personal interest as I am a second language learner of Swedish who has been exposed to South Swedish throughout.

1.1 Research questions and hypothesis

Based on the underspecification theory (Archangeli 1988), the research questions were formulated as follows: does one word accent have more specific memory traces than the other in the South Swedish dialect among native speakers (research question 1)? If the answer is yes, then which word accent has more specific memory trace (research question 2)? And why is that (research question 3)? Hence, I will test the hypotheses that in South Swedish, one word accent has a more specific memory trace than another (hypothesis 1), and that is because one word accent is default shaped by intonation and another is more lexically specified (hypothesis 2).

1.2 A snapshot of the method

This thesis analyses speech perception using an experimental method. To be specific, a neurolinguistic experiment using the EEG technique was conducted. With EEG, I looked into an event-related potential (ERP) component called *mismatch negativity* (MMN), which is a negative peak that typically occurs around 160-220ms after the stimulus onset, elicited automatically by the sudden acoustic difference in a sequence of repeated auditory stimuli. The MMN has been argued to reflect language-dependent memory traces (Näätänen et al., 1997). Politzer-Ahles et al. (2016) also provided evidence of asymmetric perception of Mandarin Chinese lexical tones using the MMN. They found tone 3 elicited a smaller MMN than other lexical tones. The present study had a similar set-up in order to find out asymmetry in the perception of South Swedish word accents and tried to interpret the asymmetry from phonological point of view.

1.3. Structure of the thesis

This thesis consists of five chapters. The first chapter gives an introduction to the topic, research question, hypothesis and method. In the second chapter, I introduce the theoretical background in Swedish word accents with a focus on South Swedish, a brief view of underspecification theory in phonology and a concise introduction of the EEG method with a focus on MMN. Chapter 3 describes the method in detail and gives a brief look of the results. Chapter 4 demonstrates the results with data analysis. Chapter 5 gives a short discussion of the results. Chapter 6 provides a concise conclusion and some suggestions for possible future research.

Chapter 2 Theoretical background

In this Chapter, I first introduce Swedish word accents with main focus on South Swedish in section 2.1. In the second section, I give a concise introduction to EEG and ERP. Section 2.3 focuses on MMN and gives a literature review of existing MMN studies on tonal patterns. At the end of this chapter, section 2.4 provides a quick look into how tonal aspects are processed in the brain.

2.1 Swedish word accents

2.1.1 A brief introduction of Swedish word accents

Tonal accents can be found in most varieties of Swedish (Finland Swedish is an exception). They may be realized as different phonetic shapes; meanwhile, the systems all consist of two different tonal configurations. Hence, the tonal accents are named *accent 1* and *accent 2*. Here is an example of the phonetic difference between accent 1 and accent 2 in Central Swedish:

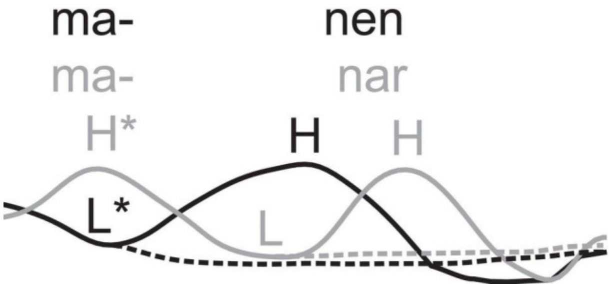

Figure 1 Pitch contour of word accents in Central Swedish with accent 1 word ^{1}*manen* (the mane) in black and accent 2 word ^{2}*manar* (manes) in grey, sourced from Roll (2022, p. 02). Solid lines represent pitch contours of the focus accented level whereas dashed lines reflect the word accented level. The two prominence levels are introduced later in this section.

The tone pattern is often regarded as the reflection of association between stem and suffix. Riad (2009, 2012, 2014) argued that lexical tones are associated with lexicalized suffixes. The stem is assigned with either a low or high tone depending on the suffix. That is to say, different suffixes can introduce different tonal patterns to the same stem, which results in different word accents. Consider the example in figure 1. The same stem *man* has accent 1 together with the singular suffix -*en* while the plural suffix -*ar* induces accent 2 onto the stem.

Some linguists argued that Swedish word accents have low functional load as determining meaning is not the main function of word accents (Elert 1964; Roll 2022). Evidence of this is the limited number of minimal pairs. In the report by Elert (1972), there are 357 minimal pairs of words that only contrast in word accent in Swedish. By saying only contrast in word accent, it means the phonetic realizations are only different in tone but not in segmental features. However, some of them are defective as minimal pairs as some minimal pairs consist of words that come from different word classes, and some minimal pairs, although they consist of words from the same word class, show morphological differences between them (Riad, 2014). Consider the example below:

^{1}and-en (the duck) - ^{2}ande-n (the spirit)

The accent 1 can be seen as the combination of a monosyllabic stem with a suffix *-en*. The suffix does not introduce any lexical tone to the stem, hence the word is pronounced in the default accent 1. The accent 2 word on the other hand, can be seen as the results of a disyllabic stem with a non-syllabic suffix, it is the stem vowel *-e* that introduces the accent 2 to the stem (Riad 1998).

However, some other linguists hold an opposite opinion. Although the minimal pairs that only contrast in word accent is not common, it does not mean that word accents are less important than segmental phonemes. Recent neurolinguistic studies on Swedish word accents suggested that word accents have their contribution in prediction, and they play an important role in an early stage of lexical retrieval in native speakers' mental lexicon (Roll 2022; Kochančikaitė, Shtyrov and Roll 2023 submitted). The relation between word accents and prediction is discussed in section 2.1.4 and related neurolinguistic studies are summarized in section 2.5.2.

In addition, it is worth mentioning that word accents have two prominence levels (Bruce 1977; Myrberg 2010; Myrberg and Riad 2015): one word accented level and one focus accented level. To summarize the two levels, I adopted a table from Riad (2014). This table demonstrates the tonal pattern in Central Swedish. The example of pitch contour with real words can be found in Figure 1 earlier in this section.

Table 1 Prominence level for word accents, sourced from Riad (2014, p.184).

Prominence level	Accent 1	Accent 2	Accent 2 in compounds	Typical functions
Focus accent	L*H	**H***LH	**H***L*H	focus, contrastive topic
Word accent	HL*	**H***L	**H***L	given material, second occurrence focus (post focally), new material (non-final in the phrase)

2.1.2 A typological view of Swedish word accents

Traditionally, Swedish dialects can be categorized into two groups by the type of accent 2: Two peaks accent 2 dialects and one peak accent 2 dialects (Gårding 1977; Riad 2006). I adapted a map with labels from Gårding and Lindblad (1973, p. 48) which clearly indicated the distribution of the two groups of dialects, shown in Figure 2. Group 0 includes dialects without tonal accents, mainly spoken by Finland Swedish speakers. Group 1 includes one peak accent 2 dialects, which can be further divided into two sub-groups (1A and 1B) within Swedish. Group 1A includes South Swedish, spoken by native people from Skåne, South Halland and Blekinge in South Sweden. Group 1B includes Dala Swedish, spoken in Dalarna and Gotland Swedish, spoken by Gotland people from the island Gotland. Group 2 involves two peak accent 2 dialects, which include Central Swedish and West Swedish. From the pitch pattern on the lower right corner of the picture we can clearly see that despite the different shape of phonetic realization, each group (except for group 0) has a clear difference between accent 1 and accent 2.

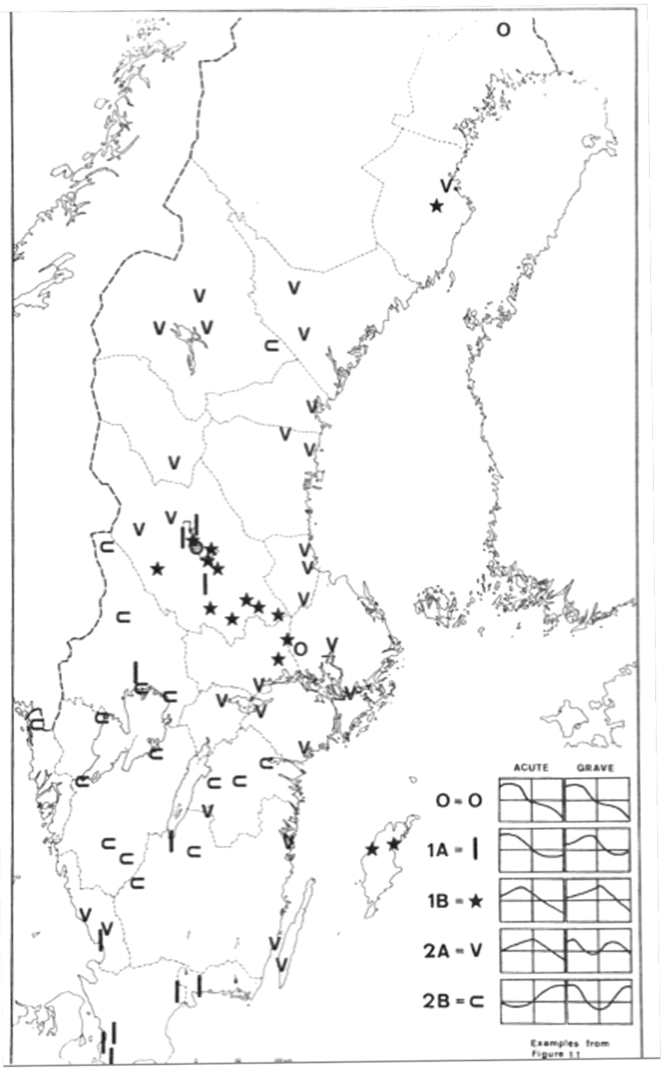

Figure 2 Geography distribution of Swedish word accents, sourced from Gårding and Lindblad (1973, p.48).

2.1.3 South Swedish word accents

First of all, I would like to point out that there are also dialectal differences within South Swedish. In the present study, however, I mainly focus on Malmö dialect, which is spoken in the biggest city in southern Sweden, as it is the most spoken (by number of speakers) variety within South Swedish.

South Swedish, as a one-peak dialect, has only one H tone in accent 2. That is to say, regardless of whether the word is in a focus or non-focus position, in a compound or not in a compound, once the tone falls, it never rises again. Another factor that is different from Central Swedish is that the tonal pattern associated with the stressed syllable (the “stem tone”) is the complete opposite in South Swedish. The tone associated with stress in South Swedish accent 1 is H, while in accent 2, it is L.

The differences in tone representation of word accents between South Swedish and Central Swedish are also shown on the sentence level. Bruce (1977) investigated Swedish word accents in sentence perspective in the Stockholm and Malmö dialects. He pointed out that the tonal patterns of word accents in focus and non-focus positions only share limited similarities. In the Malmö dialect, accent 1 words have a rising phase in the prevocalic consonant and the peak is in the stressed vowel. When the accent 1 word is in focus position, there is a steep fall beginning in the middle of the stressed vowel and ending with the post-tonic consonant. When accent 1 is in a non-focus position, on the other hand, only a mild fall is observed in the post-tonic syllable. Accent 2 words in the Malmö dialect also have a rising phase, but the peak is on the postvocalic consonant. When accent 2 is in a focus position, there is a steep fall in the post-tonic vowel. However, the falling pattern was not found in non-focus accent 2 words. In focus position, the main prosodic cue in the Malmö dialect is the fall of the F0 contour and the fall is accent independent. That is to say, accent 1 and accent 2 words are different in the timing of the fall when in focus positions. accent 1 has an earlier fall, while accent 2 has a later. Table 2 demonstrate a summarized tonal pattern of South Swedish word accent is different position.

Table 2 South Swedish word accents.

	Non-focus	Focus
Accent 1	(L)H*L	(L)H*L
Accent 2	L*HL	L*HL

2.1.4 Perception cues for Swedish word accents

Bruce and Gårding (1977) stated that the acoustic difference between accent 1 and accent 2 can be defined with the timing of the high-low contour. In South Swedish, the high-low contour in accent 2 is later than accent 1. There is a gliding rise in pitch before the falling tone in South Swedish accent 2. Ambrazaitis and Bruce (2006) indicated that both the slow raising and the late fall are important for accent 2 perception in South Swedish. Gosselke Berthelsen et al. (2022) indicated that the importance of the fall for perception of Swedish word accents is general among dialects.

2.1.5 Swedish word accents and its prediction function

For a language with a simple tonal system like several North Germanic languages, tonal structures function slightly different from those in a complex tonal language like Mandarin. Previous studies indicated that Swedish word accents are highly associated with prediction. Native speakers tend to use tonal patterns to predict word endings. That is because some tonal patterns have fewer continuations; hence, the word ending is highly predictable. On the other hand, some accents have more possible continuations, resulting in a low predictive value, which also raises the risk of prediction errors. In the case of Swedish, accent 1 is more predictively useful than accent 2 (Roll et al. 2015, 2017; Söderström 2016; Roll 2022). That is because accent 1 has fewer possible endings than accent 2, as accent 1 is associated with a well-defined set of suffixes, whereas accent 2 is associated with not only suffices but also compounds (in Central Swedish, all compounds are in accent 2, in South Swedish, some compounds are in accent 1 but most of them are also in accent 2). The predictive function or facilitative function suggests that Swedish word accents are morphophonological rather than lexical. In addition, a recent study also indicated that Swedish word accents also involved in semantic processing (Kwon 2023).

2.2 EEG and ERP

EEG is the electrical activity of the brain that can be detected by placing electrodes on the scalp, amplifying the signal, and plotting the change over time in voltage (Luck 2014, pp. 3-4). It was first discovered by Hans Berger in 1929. EEG is a common brain imaging technique in cognitive neuroscience research nowadays. There are a few reasons why EEG is more popular than others. The first advantage of EEG compared to some other brain imaging techniques like functional Magnetic Resonance Imaging (fMRI) is that it provides a more precise temporal solution as they record data over milliseconds (same for magnetoencephalography). The second advantage is that the EEG solution is non-invasive, which is both researcher-friendly and participant-friendly. In addition, EEG is an easy hands-on method, it is less expensive, and the operation is simple and does not require technicians involved. On the other hand, EEG does have some disadvantages, for example, EEG data is very easy to be polluted by movements (also true for other method such as MRI), muscle activity and magnetic interference. EEG is also limited in its source-localization quality due to the inverse problem, that is because the locations and orientations of the dipoles is uncertain. That is to say, using inverse modelling, which is the current solution, can still estimate the source of signals.

Raw EEG data, however, is coarse and noisy. To find embedded information requires isolating signals by averaging signals and time-locking signals. The time-locked, pre-processed and averaged EEG is called ERP. In linguistic research, ERP is often used to investigate the timing of language processing in the brain. The typical negative or positive peaks in ERP after receiving stimuli, which are called ERP components, are regarded as cues and evidence of different statuses of processing. Precisely, an ERP component is a scalp-recorded neural signal that is generated in a specific neuroanatomical module when a specific computational operation is performed (Luck 2014, p. 66). Over the years, numeral ERP components have been discovered. Some of them are related to visual stimuli, some are related to auditory stimuli, and others are related to semantics, etc. In the present study, I looked into an auditory-related component named MMN.

2.3 MMN

2.3.1 Introduction of MMN

MMN is an auditory ERP component discovered by Näätänen et al. (1978). It is automatically or pre-attentively elicited by an auditory stimulus differing from preceding stimuli (Näätänen et al. 1978). It is typically a negative peak that appears 160-220ms after stimulus onset, with a spatial fronto-central maximum. MMN can be evoked even when passively listening, however, strong attention and focusing may attenuate the amplitude of MMN (Woldorff et al. 1991).

MMN can be sub-categorized into two subcomponents, one is a bilateral supratemporal component located on the auditory cortex which associates with preperceptual change detection, another is a frontal component located in pre frontal cortex which associates with the initiation of involuntary attention switch caused by an auditory change (Näätänen and Kreegipuu 2011).

MMN is easy to be refracted by other early components, such as auditory N1 and visual P1. That is because MMN falls just between the N1 latency and P1 latency. Furthermore, other factors may influence the MMN, such as the intensity of sound, unexpected early tone onset, and the omission of a stimulus (Luck, 2014). Therefore, those factors should be taken into consideration when designing the experiment. Another factor that often influences the MMN is the auditory N1, which can be elicited by the pitch difference of stimuli. There is another component called identity MMN, which is exogenously evoked by the deviant that is identical to the standard. It is thought to be able to reflect pre-attentive processing of auditory stimuli. Hence, the identity MMN technique is used to minimize the influence of N1 and P1 (Pulvermüller and Shtyrov 2006, p. 54).

2.3.2 MMN and language processing

As mentioned in the last section, MMN is induced by acoustic differences. The difference is not necessarily language dependent. This is to say pure acoustic differences such as the change of frequency, duration, intensity and timbre can also elicit an MMN. However, what is the differences when the contrast in auditory stimuli is language dependent compared to language independent? Näätänen et al. (1997) did an MMN study on Finnish and Estonian vowel perception among Finnish native speakers. They found that the MMN was enhanced when the deviant was a prototype in the language compared to its non-prototype (Näätänen et al. 1997, p. 432). Based on that, Näätänen et al. (1997; 2007) argued that the MMN can reveal the existence of language-dependent memory traces. Pulvermüller et al. (2001) using MMN revealed the memory trace of words in the human brain and located the source in the left superior temporal lobe.

MMN is thought to be associated with prediction error and it can reflect the violation of memory trace of regularity. This is because the MMN reflects the comparison between a short-lived memory trace of the standards and the current stimulus. Hence, when listening to a repeated stimulus, the brain predicts the upcoming stimuli to be the same. If the new stimulus is different, it enhances the amplitude of the MMN. MMN is associated with prediction errors. This proposal has been well-received and the interpretation of the results in the present study is also based on this hypothesis.

However, when the difference between standard and deviant is language dependent, the latency of MMN can vary a lot. Zachau et al. (2005) observed that late MMN can be evoked by abstract tone patterns. The transfer of rules to long-term memory may be what causes the late MMN, which may be based on processes for extracting rules from auditory input.

2.3.3 MMN asymmetry and its relation with phonology

Sometimes, when one sound is the standard and the other is the deviant, the MMN amplitude may be higher than when it is the other way around. This asymmetry is not aroused by acoustic differences but rather the difference in the mental representation of the sound. That is to say, sounds with more specific memory induce larger MMN. In the case of Swedish word accents, one tonal accent may have less specific memory because it is underspecified, shaped by intonation by default whereas the other tonal accent has more specific memory representation because it is lexically specified.

2.4 Previous ERP studies on tonal patterns and Swedish word accents

2.4.1 Previous MMN studies

There are numeral MMN studies on lexical tone perception in Mandarin Chinese from different points of view (more to summarize). Politzer-Ahles et al. (2016) indicated that Mandarin tone 3 has more surface realizations compared to other lexical tones, and it is the tone that has the most tone sandhi in continuous speech. Based on these facts, they argued that Mandarin tone 3 has less specific memory trace than other tones. Politzer-Ahles et al. provided evidence from MMN experiments, and their results showed that tone 3 elicited a relatively smaller MMN. Kann et al. (2008) found that high-rising tones elicited left-lateralized late MMN. Zora et al. (2016) investigated lexical specification in stress pattern in Swedish using MMN and they discovered that MMN amplitude was greater for the phonologically stressed word than for the lexically stressed word. Kochančikaitė, Shtyrov and Roll (2023 submitted) were the first-ever researchers using MMN to investigate Swedish word accents. They compared lexical MMN and syntactic MMN enhanced by different word accents and found out that tonal pattern in Swedish is useful for lexical retrieval in an early stage of word activation in the brain. They also argued that although Swedish word accents have low functional load as there are only limited minimal pairs, tonal accents are still an integral part of words in native speaker's brain's mental lexicon.

2.4.2 Other ERP studies on Swedish word accents

As mentioned in section 2.1.4, Swedish word accents are believed to contribute more on prediction. Throughout the years, neurolinguistic studies have provided ample evidence for the prediction function of Swedish word accents at the brain level. When coming to the neural level, prediction maps to pre-activation in the brain. Accent 1 in Swedish, which is believed to be more predictively useful, produces stronger pre-activation. The pre-activation can be detected from the scalp. In 2010s, a negative trending ERP component appeared around 100ms after tone onset named. It was later named the *pre-activation negativity* (PrAN) and thought to be the evidence of pre-activation (Roll et al. 2015, 2017; Söderström et al. 2016). Highly predictive accent 1 elicited more significant PrAN than accent 2, implying that PrAN correlates with predictive value. A study on South Swedish found that PrAN component for South Swedish accent 1 shows up slightly later than Central Swedish, at 220-280ms after fundamental frequency (F0) onset. Hence, it was thought to support the hypothesis suggesting that PrAN is not the reflection of an acoustic difference of word accents (Roll 2015, p.159). There is also evidence showing that PrAN can be elicited even when listening passively. However, the latency was around 400 - 600ms later than non-passive listening (Kochančikaitė, Shtyrov and Roll. 2023 submitted). On the contrary, MMN was thought to be able to reflect prediction error. The word accent that arouses increased prediction error should induce larger MMN.

2.5 Tone processing in the brain

Pure acoustic features like musical notes and low-level linguistic elements such as intonation are mainly processed in the right hemisphere of the brain (Zatorre and Gandour 2008). That is because the right hemisphere is better at slow integration and spectral processing (Poeppel 2003). Lexical tone perception in the brain, on the contrary, has been observed to show an asymmetry favoring the left hemisphere due to the myelination of the cortices' surface. Previous studies show that in L1 speakers of languages with both complex and simple tonal systems recruit the left secondary auditory cortex (planum temporale) while processing word tones (Xu et al. 2006; Roll et al. 2015; Söderström et al. 2017; Liang et al. 2018; Schremm et al. 2018 idea for accent storage). For Swedish word accents, Roll et al. (2015) provided functional data of left planum temporale, Herschel's gyrus and superior temporal gyrus in Swedish word accent processing. Studies also indicated that cortical thickness of left PT and pars opercularis (Brodmann area 44 or BA44) are related to Swedish word accent processing among native speakers (Roll et al. 2015; Schremm et al. 2018; Novén et al. 2019, 2021). Liang et al. (2018) conducted a meta-analysis and argued that "only L1 tonal language speakers consistently recruited the left superior temporal gyrus in lexical tone perception, supporting that language experience shapes lexical tone as a phonetic feature in defining lexical meaning". Feng et al. (2018) found that the left inferior parietal lobule and the left superior temporal gyrus were involved in the pitch perception of the tone through representational similarity analysis to complete the speech representation and classification processing.

For Swedish word accents, since different word accents are related to different values of predictability, the brain also has different ways to process them. Söderström et al. (2017) indicated that activity in the left inferior parietal area is triggered by more predictive stems with fewer potential continuations.

The results of previous studies suggest that left temporal region should be particularly inclined to show results in the present study.

Chapter 3 Method

In this chapter I describe the design and procedure of the MMN experiment in detail. Section 3.1 introduces the experiment design. Section 3.2 gives information about the stimuli, including how they were recorded, selected, and manipulated. In section 3.3 I describe the EEG recording procedure. How the raw EEG data was processed is explained in section 3.4. Section 3.5 describes the statistic method used for data analysis and section 3.6 explains the procedure of source localization.

3.1 Experiment design

In general, the experiment runs under the passive oddball paradigm (Squires et al. 1975). In a typical oddball design, there are repeated sequential stimuli named *standards,* while the interrupting stimuli are called *deviants*. Only auditory stimuli from recordings were used in the experiment. The stimuli consist of a minimal pair of words that only contrasted in word accent: ^{1}*kullen* (the litter) in accent 1 and ^{2}*kullen* (the hill) in accent 2. They functioned as each other's standard and deviant. The experiment included two blocks. In the first block, the accent 1 word was standard while the accent 2 word was deviant. In the second block, however, the accent 2 word switched to standard, whereas the accent 1 word functioned as deviant. Each block had 866 valid stimuli, of which 736 were standards while the other 130 stimuli were deviants (15% deviants). The order of the stimuli was pseudorandomized, and deviants were separated by at least one standard.

There were also 150 extra stimuli that were the same as the standards but not taken into the analysis. They were the first 20 stimuli and the first standard after each deviant. The first 20 stimuli of each block were designed to give participants a warm-up. Hence, they were discarded in the analysis. The first standards after each deviant were also discarded. The purpose was to avoid the MMN caused by the preceding deviant. Hence, there were 2,032 stimuli in total to which each participant was exposed, 300 of which were not used in the analysis. Stimulus onset asynchrony (SOA) was randomly shuffled around 1,300 ± 50ms in 10ms step. The role of the SOA variation was to minimalize the effect of alpha waves caused by regularly repeated stimuli (Luck 2014). There was also a three-minute break between each block for participants to rest a little bit and for the researcher to check the impedance again.

Table 3 Experiment design.

	Condition	Standard → deviant	SOA
Block 1	Accent 2 as mismatch	^{1}*kullen* (85%) → ^{2}*kullen* (15%)	1,300 ± 50ms
Break			
Block 2	Accent 1 as mismatch	^{2}*kullen* (85%) → ^{1}*kullen* (15%)	1,300 ± 50ms

3.2 Stimuli recording and selection

A male speaker (age = 23) from Malmö recorded the stimuli for the experiment. All recordings were done in the anechoic chamber in Humanities Lab of Lund University. Sounds were recorded as mono sound with a sampling frequency of 44,100Hz in 24-bit sample format in Audacity. The speaker was asked to produce the two words ^{1}*kullen* and ^{2}*kullen* several times for the researcher to select the best realizations. The reason why this minimal pair was chosen was because both words are composed of a noun stem and a definite article suffix, and the first syllable is the same, and the following consonant is sonorant which has less effect on the pitch contour than obstruent. After recording, all sounds were saved as .wav files with 16-bit PCM audio quality and sent to Praat (Boersma 2001) for manipulation.

In Praat all sounds were normalized in amplitude using the plugin Praat Vocal Toolkit (Corretge 2012). Sounds with flaw (e.g., non-modal voice quality in the first syllable, unclear articulation, wrong pronunciation etc.) were discarded. Finally, only one accent 1 word and one accent 2 word were selected as stimuli. Both of them had clear pitch curves extracted by Praat. Their F0 onset was 89ms after word onset. Figure 3 demonstrates the pitch contour of those two sounds in semitone referenced to 100Hz. Pitch curves in the figure are smoothed using a 10Hz bandwidth filter. The gap in the pitch curve is caused by creakiness, which is considered to be a common but not default voice quality for accent 1 words in South Swedish (Hjortdal 2022). However, my speaker produced creaky voice in both accent 1 and accent 2 words.

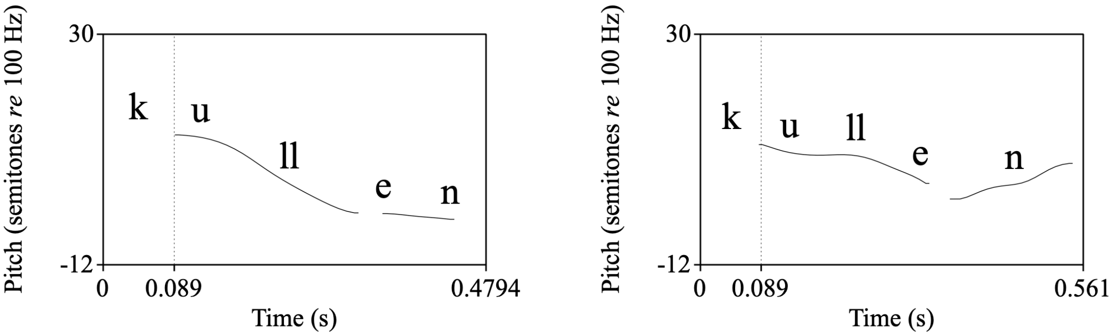

Figure 3 Pitch curve of two stimuli words, with ^{1}kullen on the left and ^{2}kullen on the right.

3.3 EEG recording

3.3.1 Participants

There were 19 South Swedish speakers from region Skåne recruited in the experiment. Twelve of them were male and the other seven of them were female. Their age range is from 19 to 37 years old, with a mean of 24 and a standard deviation (SD) of 4.04. Prior to the experiment, they all carefully read and signed a consent form, answered a questionnaire about language background, and did the Edinburgh Handedness test (Oldfield 1971). According to their answers to the questionnaire and the text, they were all South Swedish L1 speakers who grew up in Skåne or mainly in Skåne, right-handed (Their Laterality Quotients were all above 40, with mean of 90.47 and SD of 11.67, detailed information can be found in Appendix) and recorded without any neurological disorder. They were all compensated by a digital movie ticket after completing the whole experiment.

3.3.2 Recording procedure

The recording took place in the Humanities Lab of Lund University. 64-channel EEG caps with Ag/AgCl electrodes in EasyCap 10-20 system were used for recording. The distribution of electrodes is shown in Figure 4. Two individual mastoid electrodes (placed on mastoid bone by both earlobes) were used as reference electrodes and four ocular electrodes (two vertical placed above and below left eye and two horizontal placed next to both canthi) were used to measure electrooculography. Scalp impedances were reduced to under 5kΩ before recording by using abrasive electrolyte gel. Sampling rate of EEG recording was 1,000Hz. In the recording configuration, AFz was served as ground and mastoids were served as online references. Time locking points which were the F0 onset of each stimulus word were sent to the recording computer as triggers via a parallel port by PsychoPy2 (Peirce 2007, 2009). During the EEG recording, all participants sat still on a comfortable armchair and watched a silent movie playing in a small fixed screen right in front of them. They were informed to concentrate on the movie while passively listening to the sounds playing in the headphone (The headphone used in the experiment was KOSS SB/45). The recording took approximately two hours in total.

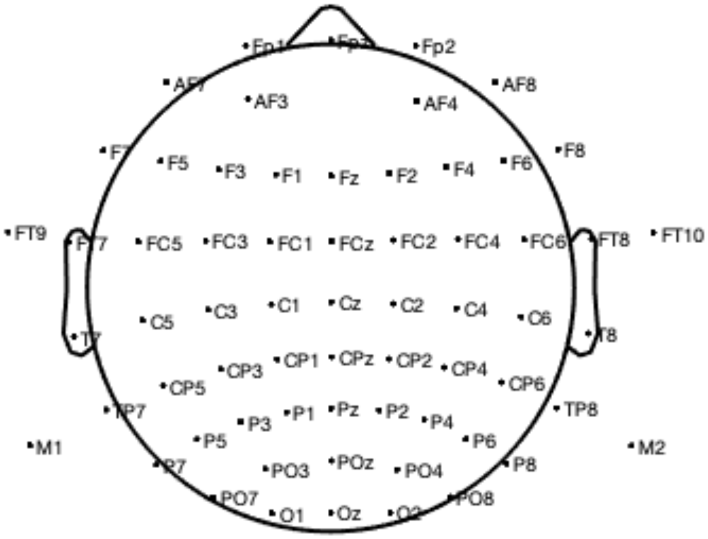

Figure 4 Channel locations (without ground channel AFz and ocular electrodes).

3.4 Data processing

After recording, the data was analysed using EEGLAB (v2021.1) (Delorme and Makeig, 2004) in MATLAB. First, I looked through the raw data and reject irrelevant phases and removed bad channels if they existed. Bad or missing channels were interpolated (spherical spline interpolation) and no more than one channel within individual participants was interpolated. Then, the data was re-referenced to average reference. After that, a bandpass filter with cut-off frequencies 0.05 Hz and 30 Hz was applied to remove low and high frequency noise. Independent component analysis (ICA) was applied to identify blinks, movements and channel noise. The algorithm was the default one named *runica* in EEGLAB. After running ICA, I rejected several ICA components (including channel noise, ocular movements and muscle activities) with the assistance of ICA labelling. The number of rejected components varied from 2 to 10 among participants (I tried to keep the number as low as possible, 10 is the extreme number from one participant). Target epochs were extracted from the data with an 800ms time window from -200ms to 600ms after the time lock point. Any trial where the signal exceeded ±100μV was also rejected. After processing, there were 32,332 trails in total marked as good, including 2,439 trails of accent 2 deviant and 13,676 trails of accent 1 standard from block 1, 2,454 trials of accent 1 deviant and 13,763 trails of accent 2 standard from block 2. Averaged EEG were plotted by condition against each other.

3.5 Statistical analysis

Cluster-based permutation analysis was performed using the Fieldtrip toolbox (Oostenveld 2011) to detect if there is any significant difference between the amplitude of standard and deviant. The reason I chose this statistical method is that the sample sizes of standard and deviant are not symmetric. Permutation analysis is relatively straightforward, non-parametric and gives unbiased scalp distribution. Permutation test analyses a 15ms time window around each peak. Since it is a one-tailed test, significant threshold (α) was set to 0.025.

3.6 Source localization

Source localization was conducted in Brainstorm (Tadel et al. 2011) using the default anatomy. Noise covariance matrix was derived from recording. Head model was computed using OpenMEEG BEM (Gramfort et al. 2010; Kybic et al. 2005). Finally, source was estimated using minimum-norm imaging method.

Chapter 4 Results and statistics

In this chapter, I first present the results by giving average plotting in section 4.1. Then I analyse the data using permutation analysis for inferential statistics, show the results in topographies and describe the MMN asymmetry in section 4.2. Section 4.3 displays estimated source.

4.1 Results plotting

Average EEG plotting from each block extracted from FC3 electrode is shown in Figure 5. I chose FC3 for presentation because it appears in every significant negative cluster in the permutation analysis, which will be presented in the next section. From the plot, we can already see the difference between blocks. In block 1, the accent 2 deviant aroused a negative trending wave after around 200ms, and the negativity spread out until around 400ms. The negative cluster had two peaks, a smaller one around 250ms and a more amplified one around 350ms. In block 2, an early negative cluster was detected for the accent 1 deviant around 170ms and another negative trending wave around 350ms. It is noticeable that there is a negative going wave after 500ms in the difference curve, which is not found in the other block.

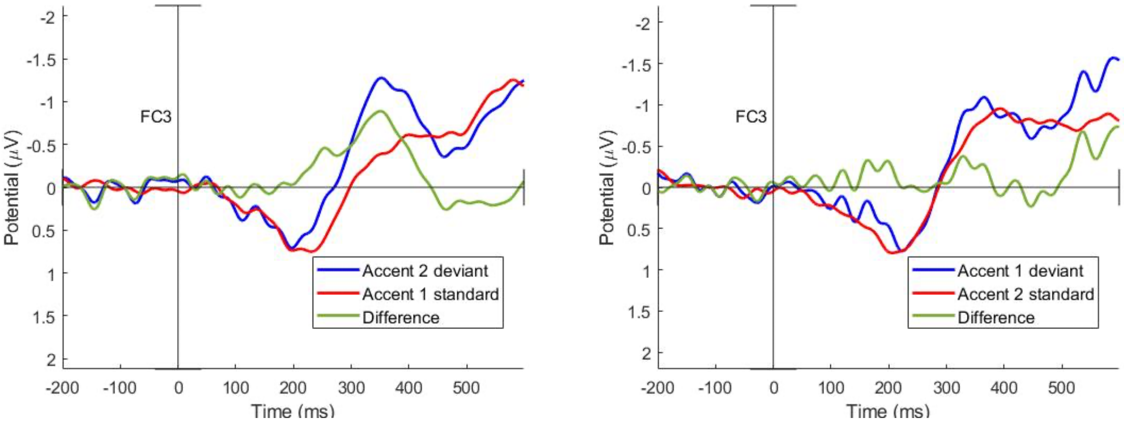

Figure 5 Average ERP of 19 participants at FC3 electrode with block 1 on the left and block 2 on the right.

Next, I subtracted the identity MMNs, which are accent 1 deviants compare to accent 1 standards and accent 2 deviants compare to accent 2 standards. This was to minimize the acoustic effect. Figure 6 demonstrates identity MMN plots at the FC3 electrodes. It is clear that both accents elicited a negative trending wave at the latency range from 300 to 400ms. However, before that, both accents arouse an earlier negative wave, with accent 1 even earlier around 150 - 200ms and accent 2 slightly later around 250ms.

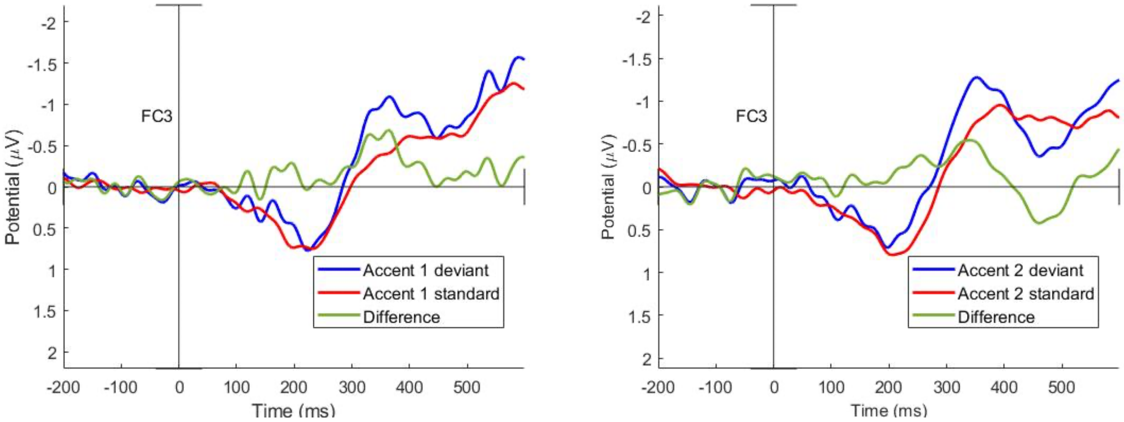

Figure 6 Identity MMN of 19 participants at FC3 electrode with accent 1 on the left and accent 2 on the right.

Then I plotted all deviants against all standards. The plot from channel FC3 is shown in Figure 7. This is to reduce the acoustic effect. Although this will make the MMN not constrained to a specific word accent, it still reflects the general word accent effect. From the plot, it is likely to say the non-acoustic word accent processing happens around 310-380ms.

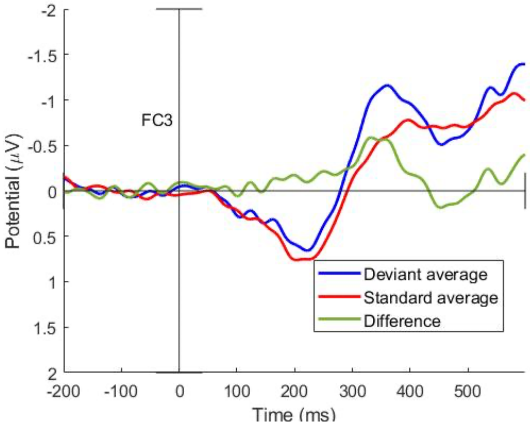

Figure 7 Average deviant vs. average standard at FC3 electrode.

4.2 Results of statistical analysis

In block 1 where accent 2 was the deviant, there was no early MMN detected, but a significant large late MMN extending from 250 to 400ms after the time locking point can be observed. The late negative trending wave had two peaks, one around 246-261ms at F3, Fz, FC3, FC4, C3, C4, CP3, CPz, CP4, F5, F1, F2, FC5, FC1, FC2, C5, C1, C2, CP1 and CP2 electrodes (p = 0.003) and another around 342-357ms at F3, Fz, F4, FC3, FC4, C3, C4, CP3, CP4, AF3, F5, F1, F2, FC5, FC1, FC2, C5 and C2 electrodes (p = 0.002).

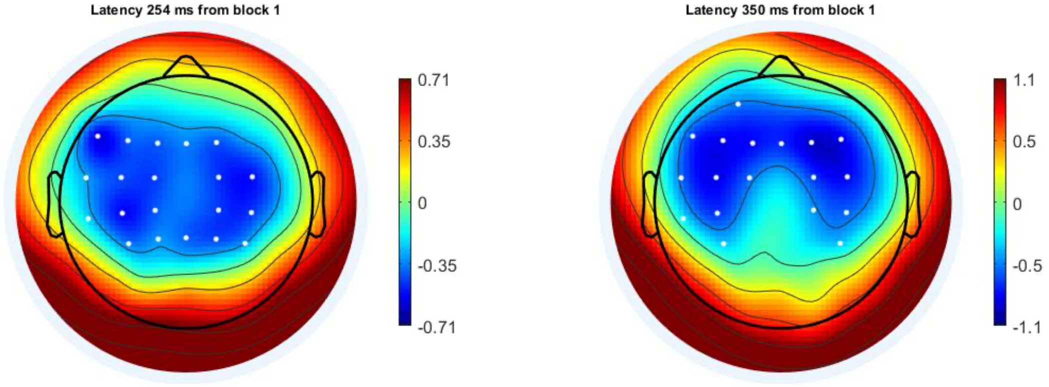

Figure 8 Topographies of significant electrodes distribution, with first peak at 254ms on the left and second peak at 350ms on the right.

When we focus on block 2, in which accent 1 was deviant. An early MMN was observed around 160-175ms after time locking point. A significant negative cluster (p = 0.019) was detected at F3, Fz, FC3, FCz, Cz, F1, F2, FC1 and FC2 electrodes. There was also an apparent effect around 190 - 205ms after time lock point at C3, CP3, FC5 and C5 electrodes but it was not significant (p = 0.0969). No significant result was found for the late MMN either in this block, although left anterior electrodes show some tendency of negative peak (e.g., a small negative peak can be seen in the right plot around 300 - 350ms in Figure 4). However, there was a significant negative wave in frontal area at around 592ms (585-600ms), including channels F3, Fz, F4, FC3, FCz, FC4, F1, F2, F6, FC5, FC1, FC2, FC6 and C2 (p = 0.011). Because of the extreme late latency, it is possible to assume that this negative wave might not be an MMN component. Hence, this negative going wave was not taken into consideration when drawing conclusion. However, what may elicit this late component is discussed in the next chapter.

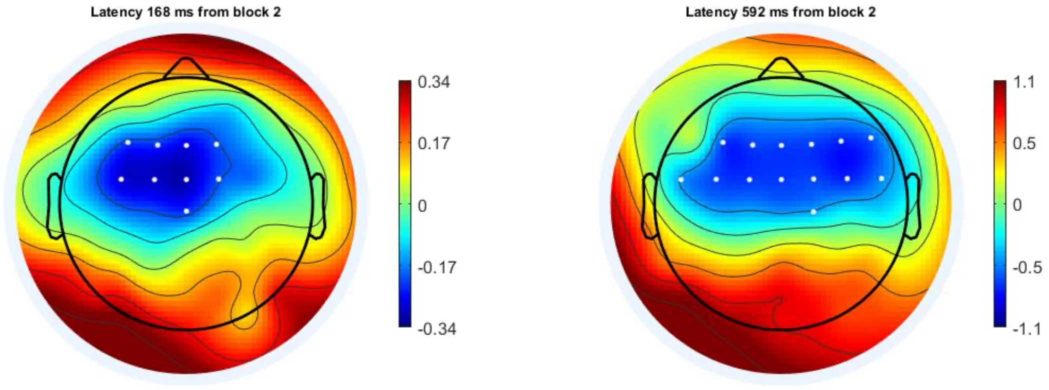

Figure 9 Topography of significant electrodes distribution, with first peak at 168ms on the left and second peak at 592ms on the right.

Moving on to the identity MMN, for the accent 1 identity MMN, no significant early cluster been detected. However, a significant cluster was detected after 300ms, with the first peak around 323 - 338ms at FC3, FC4, C3, C4, CP3, CPz, CP4, Pz, FC1, C6, CP1, CP2 and P1 electrodes (p = 0.008), and the second peak happens around 354 - 369ms at F3, Fz, FC3, FC4, C3, C4, CP3, F1, FC5, FC1, FC2, C5 and C6 electrodes (p = 0.015).

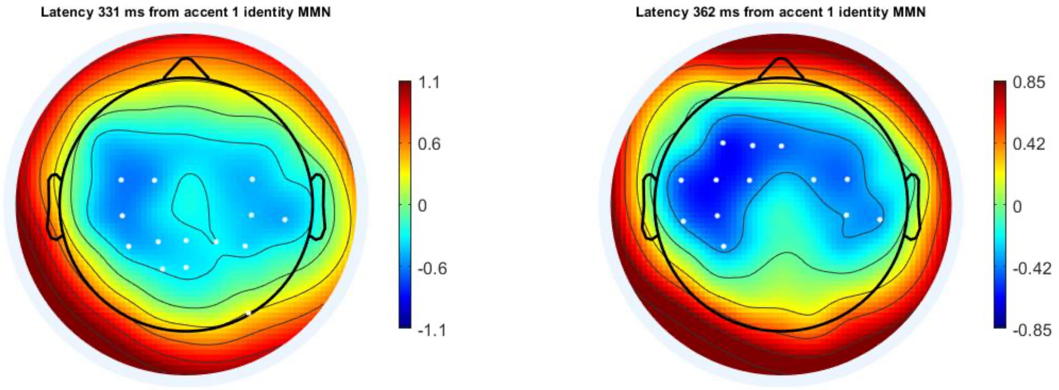

Figure 10 Topography of significant electrodes distribution at 331ms on the left and 362ms on the right from accent 1 identity MMN.

For the accent 2 identity MMN, no significant negative cluster was detected at 246 - 261ms after the time locking point.

Finally for the average deviants vs. average standards MMN, a significant negative cluster from 310 to 380ms including F3, FC3, C3, CP3, F5, F1, FC5, FC1 and CP1 was found (p = 0.015). Because it reflects how word accent are processed, I will use the word 'accent MMN' to describe this negative going component. Figure 2 shows the distribution of significant electrodes.

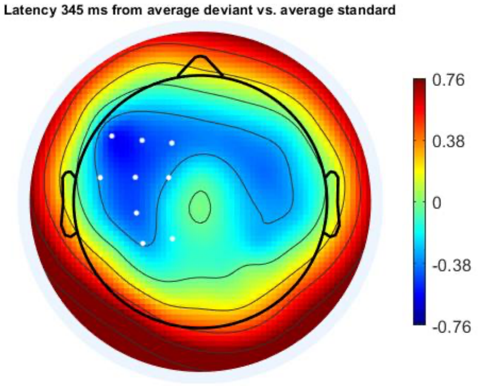

Figure 11 Topography of significant electrodes distribution of accent MMN at 345ms.

4.3 Results of source localization

Figure 12 shows the estimated source of the general accent MMN at latency 345ms. The anterior temporal lobe, namely BA38 on the left side is the most activated part, suggesting that the word accents play a role in semantic processing in general.

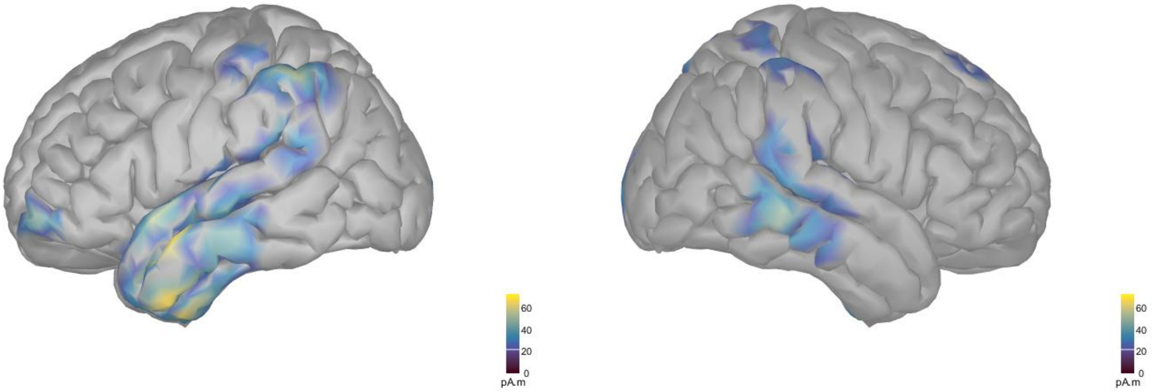

Figure 12 Source estimation of accent MMN at 345ms.

Figure 13 displays the estimated source of the accent 1 vs. accent 2 MMN at 168ms. The signals mainly come from bilateral superior temporal gyrus, superior temporal sulcus and middle temporal gyrus, indicating lower stage of speech processing. BA44/45 areas on the right show strong activation, suggesting that the main process at this time window is prosodic integration, mainly intonation (Hesling et al. 2005; Wildgruber et al. 2005). Other areas that showed strong activation include BA21/22 on the left, suggesting that semantic processing is involved in the perception of accent 1, and BA38 on the right, which is related to irony processing (Wakusawa et al. 2007) and identification of familiar voice (Nakamura et al. 2001).

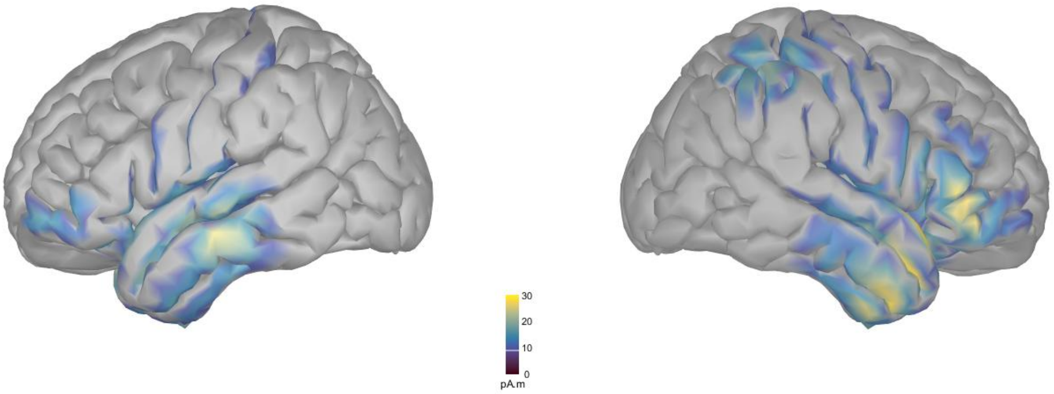

Figure 13 Source estimation of accent 1 vs. accent 2 MMN at 168ms.

The sources of the accent 2 vs. accent 1 MMN at 350ms is shown in Figure 14. From the figure it is noticeable that temporal lobes on both sides are activated, meanwhile right temporal lobe shows slightly more stirring than left one. Comparing with the general accent MMN at 345ms, which shows a similar source, it implies that the accent 2 vs. accent 1 MMN at around 350ms is more likely induced by phonological difference rather than acoustic difference. The highest pA.m level is detected in the BA21 area on the right, suggesting that this MMN is likely to be evoked by prosodic integration (Hesling et al. 2005; Ethofer et al. 2006).

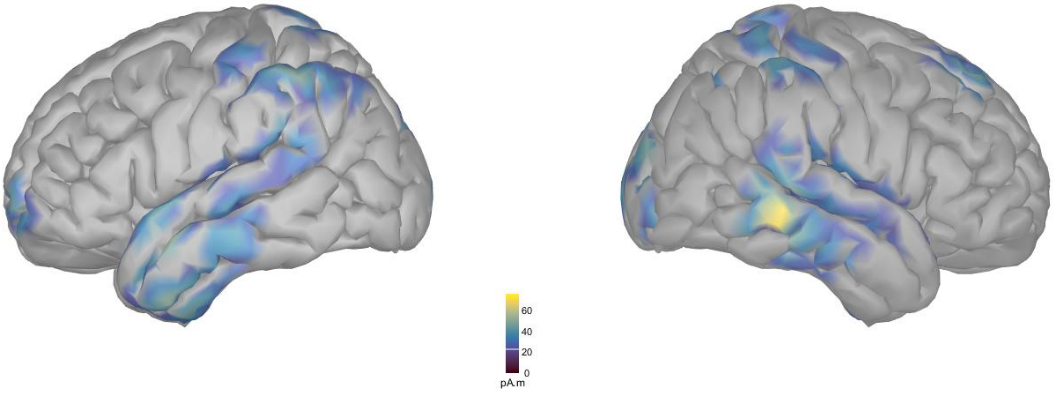

Figure 14 Source estimation of accent 2 vs. accent 1 MMN at 350ms.

Chapter 5 Discussion

In this chapter I summarize the asymmetry in MMN in section 5.1. In section 5.2 I try to identify the late component in standard vs. standard plot. Section 5.3 gives an evaluation of method.

5.1 Summarizing the asymmetry observed from the MMN results

From the current results, it seems that accent 1 elicits significant early MMN while accent 2 does not. The situation is reversed in late MMN, where accent 2 evokes a larger MMN and accent 1 does not. Hence, the MMN aroused by the different word accents is not symmetric in temporal domain. The asymmetry can also be found in amplitude dimension, as accent 2 elicited more amplified MMN. There is also asymmetry in the distribution of significant MMN recorded around the scalp. The MMN induced by accent 1 has a clear left anterior lateralization, with the maximum at FCz, FC1 and FC3. Accent 2 MMN on the other hand, does not show a clear lateralization. It has more significant electrodes on the left side but the amplitude is larger on the right hemisphere.

However, the remaining question is what factors induce the MMNs. For the early MMN elicited by accent 1, we could deduce that the MMN is not induced by phonological but rather pure acoustic differences given the fact that accent 1 enhanced early MMN when an accent 2 word functioned as standard but no significant early MMN when an accent 1 word was standard. The late accent 1 identity MMNs, on the other hand, did not have any corresponding accent 1 vs. accent 2 MMN. The reason could be that accent 1 has more predictive value. That is to say, when hearing the tone onset, the remaining part of the word is highly predictable, resulting in low prediction error. Hence, no significant late MMN is detected. The absence of late accent 1 vs. accent 2 MMN also indicates that accent 1 and accent 2 have similar representations in the mental phonology. Native speakers differentiate accent 1 from accent 2 mainly relying on acoustic differences rather than phonological differences.

Table 4 MMN comparison when accent 1 as deviant.

Latency	Standard	Significant channels	p value
160-175ms	accent 2	F3, Fz, FC3, FCz, Cz, F1, F2, FC1, FC2	0.019
	accent 1	/	
323-338ms	accent 2	/	
	accent 1	FC3, FC4, C3, C4, CP3, CPz, CP4, Pz, FC1, C6, CP1, CP2, P1	0.008
354-369ms	accent 2	/	

	accent 1	F3, Fz, FC3, FC4, C3, C4, CP3, F1, FC5, FC1, FC2, C5, C6	0.015

When accent 2 was the deviant, no early MMN was detected, but a significantly negative trending wave was observed after 200ms with two notable peaks. These prominent MMNs do not have corresponding identity MMNs in the same time window. However, the accent MMN shows a similar pattern in both amplitude, distribution and estimated source. Hence, this late component is more likely to be phonologically driven. One possible explanation for the late accent 2 vs. accent 1 MMN is that accent 2 is more lexicalized. The late MMN is similar in the temporal domain to late MMNs elicited by lexical tones in a previous study in Mandarin lexical tone (Politzer-Ahles et al. 2016). Despite the fact that left lateralization is a common finding for MMNs caused by lexical tones in previous studies, the late MMN induced by accent 2 does not show clear left lateralization. Another possible reason why the MMN peak is later for accent 2 is markedness. As discussed in Chapter 2, some linguists consider accent 2 to be marked, suggesting that it is encoded with more specific linguistic information, which in turn requires longer time to process in the brain. This longer processing time results in the late MMN observed in this experiment. The late accent 2 vs. accent 1 MMN, particularly the second peak after 300ms, can be deemed as the evidence showing that accent 2 may have more language dependent memory in the mental lexicon.

Table 5 MMN comparison when accent 2 as deviant.

Latency	Standard	Significant channels	p value
246-261ms	accent 1	F3, Fz, FC3, FC4, C3, C4, CP3, CPz, CP4, F5, F1, F2, FC5, FC1, FC2, C5, C1, C2, CP1, CP2	0.003
	accent 2	/	
342-357ms	accent 1	F3, Fz, F4, FC3, FC4, C3, C4, CP3, CP4, AF3, F5, F1, F2, FC5, FC1, FC2, C5, C2	0.002
	accent 2	/	

Amplitude difference can also be found if we compare the significant MMN evoked by accent 1 and accent 2 regardless of the temporal discrepancy. Based on previous studies, two explanations can be posed. One possible explanation is that accent 2 may have more specific memory traces than accent 1. The source localization also shows that the left anterior temporal lobe and left superior temporal lobe is the main contributor to the late MMN, which is similar to the MMN evoked by individual word memory trace in the study conducted by Pulvermüller et al. (2001). Hence, it suggests that the processing of accent 2 to some extents plays an important role in word selection at the early stage, showing that accent 2 is probably an integral part of words in the mental lexicon of native speakers. That is to say, accent 2 is more lexically specified than accent 1. Another explanation could be that accent 1 is more constraining than accent 2, as accent 1 words have less possible upcoming suffixes, which lowers the possibility of prediction error. On the other hand, accent 2 elicited larger prediction error, resulting in a larger MMN.

Table 6 Peak and amplitude comparison.

	Early peak	Late peak	Most amplified electrode(s)
accent 1 MMN	+	-	FCz, FC1, FC3
accent 2 MMN	-	++	FC5 (first peak) / FC2, FC4 (second peak)
accent 1 identity MMN	-	++	F3, C3 (first peak) / FC3 (second peak)
accent 2 identity MMN	-	-	

5.2 Identifying the late component in block 2

As mentioned in section 4.2, there is a negative going wave in block 2 at the latency after 500ms. Trying to identify the component, accent 1 standard vs accent 2 standard were plotted against each other, Figure 16 shows the plot at FC3 channel. Cluster based permutation analyses were carried out with two time windows, first is 360-410ms after tone onset, second was 500-600ms after tone onset. Results show that accent 2 elicited a significantly more negative wave compared to accent 1 from 360-410ms (F3, Fz, F4, FC3, FCz, FC4, C3, Cz, C4, AF3, AF4, F1, F2, F6, FC5, FC1, FC2, C5, C1 and C2, p = 0.002), whereas accent 1 elicited a more negative wave after 500ms (FC3, FCz, C3, Cz, FC1, C1 and C2, p = 0.049). The later negative wave could probably be regarded as the influence of a delayed PrAN, which was observed in a previous MMN study as well (Kochančikaitė, Shtyrov and Roll 2023 submitted). The early negative wave induced by accent 2, however, does not have any similarity with results in previous ERP studies on Swedish word accents, there is also no convincible explanation to be made from the present study on what it is.

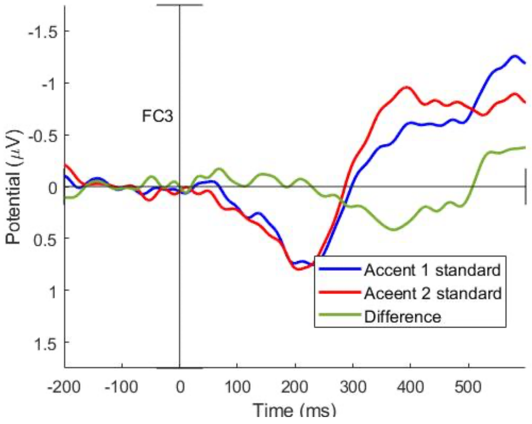

Figure 15 Accent 1 standard vs. accent 2 standard.

5.3 Evaluation of the method

The experiment runs under a passive oddball paradigm, in which participants only passively listen to the sound playing in the headphone without any additional task. This means the experiment could be fairly boring and the participants might feel sleepy during the recording. Sleepiness and boredom might lead to alpha wave production. In fact, several participants produced large alpha waves during the second half of the recording. Although ICA analysis could identify some of them, some of them still remained in the dataset.

Another problem is the word selection. The accent 1 word chosen in this experiment ^{1}*kullen* is less frequently used than the accent 2 word ^{2}*kullen*. Hence, it requires longer time in the first stage of word processing, which is retrieval the word from mental lexicon. This factor might be also contributed to the temporal asymmetry.

Chapter 6 Conclusion

In conclusion, asymmetry was observed from MMN data in both temporal, amplitude and source domains. Temporal-wise, Accent 1 enhanced an early acoustic-driven MMN while accent 2 enhanced a late phonological MMN. Amplitude-wise, accent 1 induced lower MMN than accent 2, which supports the hypothesis that accent 2 has a more specific memory. That could be because accent 2 is more marked and lexicalized and accent 1 is shaped by default, which makes accent 2 more specified in the phonological representation in the brain, or simply because accent 2 cues more possible continuations, which makes it more complicated to process.

Another exploration is that the perception of both word accents recruited cortical areas in the ventral stream of speech perception (Hickok and Poeppel 2007), suggesting that they all facilitate the procedure of lexical interface and combinatorial processing, i.e., mapping sound to semantic representations, and constructing the integration of grammatical and semantic structures. However, due to the limitation of time, this study did not conduct a statistical analysis of the results of source localization, which imposed restrictions on the interpretation of MMN sources. It will be better to apply a z-test to reanalyse the estimated source in the future.

References

Ambrazaitis, G., & Bruce, G. (2006). Perception of South Swedish word accents. *Working Papers 52, Proceedings from Fonetik*, 5-8.

Archangeli, D., 1988. Aspects of underspecification theory. *Phonology*, *5*(2), pp.183-207.

Boersma, P., 2001. Praat, a system for doing phonetics by computer. *Glot. Int.*, *5*(9), pp.341-345.

Bruce, G., 1973. Tonal accent rules for compound stressed words in the Malmö dialect. *Working papers/Lund University, Department of Linguistics and Phonetics*, *7*.

Bruce, G., 1977. Swedish accents in sentence perspective. *Working papers/Lund University, Department of Linguistics and Phonetics*, *12*.

Bruce, G. and Gårding, E., 1978. A prosodic typology for Swedish dialects. In *Nordic prosody: papers from a symposium* (pp. 219-228). Department of Linguistics, Lund University.

Corretge, R., 2012. Praat vocal toolkit: A praat plugin with automated scripts for voice processing. *Software package (http://www. praatvocaltoolkit. com)*.

Delorme, A. and Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, *134*(1), pp.9-21.

Elert, C.C., 1964. *Phonologic studies of quantity in Swedish: based on material from Stockholm speakers* (Doctoral dissertation, Stockholms kommunalförvaltning).

Elert, C.C., 1972. Tonality in Swedish: Rules and a list of minimal pairs. In: Firchow, E. ed. *Studies for Einar Haugen: Presented by Friends and Colleagues*. Berlin, Boston: De Gruyter Mouton, pp. 151-173.

Ethofer, T., Anders, S., Erb, M., Herbert, C., Wiethoff, S., Kissler, J., Grodd, W. and Wildgruber, D., 2006. Cerebral pathways in processing of affective prosody: a dynamic causal modeling study. *Neuroimage*, *30*(2), pp.580-587.

Feng, G., Ingvalson, E.M., Grieco-Calub, T.M., Roberts, M.Y., Ryan, M.E., Birmingham, P., Burrowes, D., Young, N.M. and Wong, P.C., 2018. Neural preservation underlies speech improvement from auditory deprivation in young cochlear implant recipients. *Proceedings of the National Academy of Sciences*, *115*(5), pp.E1022-E1031.

Felder, V., Jönsson-Steiner, E., Eulitz, C. and Lahiri, A., 2009. Asymmetric processing of lexical tonal contrast in Swedish. *Attention, Perception, & Psychophysics*, *71*, pp.1890-1899.

Gårding, E. and Lindblad, P., 1973. Constancy and variation tn Swedish word accent patterns. *Working papers/Lund University, Department of Linguistics and Phonetics*, *7*.

Gårding, E., 1977. The Scandinavian word accents. *Working papers/Lund University, Department of Linguistics and Phonetics*, *8*.

Gosselke Berthelsen, S., Horne, M., Shtyrov, Y. and Roll, M., 2022. Native language experience shapes pre-attentive foreign tone processing and guides rapid memory trace build-up: An ERP study. *Psychophysiology*, *59*(8), p.e14042.

Gramfort, A., Papadopoulo, T., Olivi, E. and Clerc, M., 2010. OpenMEEG: opensource software for quasistatic bioelectromagnetics. *Biomedical engineering online*, *9*, pp.1-20.

Hesling, I., Clément, S., Bordessoules, M. and Allard, M., 2005. Cerebral mechanisms of prosodic integration: evidence from connected speech. *Neuroimage*, *24*(4), pp.937-947.

Hickok, G. and Poeppel, D., 2007. The cortical organization of speech processing. *Nature reviews neuroscience*, *8*(5), pp.393-402.

Hjortdal, A., 2022, June. Creaky voice in South Swedish accent 1. In *Proceedings of Fonetik 2022: Fonetik 2022-the XXXIIIrd Swedish Phonetics Conference*. Department of Linguistics, Stockholm University.

Kaan, E., Barkley, C.M., Bao, M. and Wayland, R., 2008. Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese: An event-related potentials training study. *BMC neuroscience*, *9*, pp.1-17.

Kiparsky, P., 1985. Some consequences of lexical phonology. *Phonology*, *2*, pp.85-138.

Kochančikaitė, R., Shtyrov, Y. and Roll. M., 2023. *Neurophysiological evidence of rapid tone-based lexical access*, (submitted).

Kwon, J., 2023. *The role of word accents in semantic processing in South Swedish—An ERP study with minimal pairs* (Master Thesis, Lund University).

Kybic, J., Clerc, M., Abboud, T., Faugeras, O., Keriven, R. and Papadopoulo, T., 2005. A common formalism for the integral formulations of the forward EEG problem. *IEEE transactions on medical imaging*, *24*(1), pp.12-28.

Lahiri, A., Wetterlin, A. and Jönsson-Steiner, E., 2005. Lexical specification of tone in North Germanic. *Nordic journal of linguistics*, *28*(1), pp.61-96.

Liang, B. and Du, Y., 2018. The functional neuroanatomy of lexical tone perception: an activation likelihood estimation meta-analysis. *Frontiers in neuroscience*, *12*, p.495.

Luck, S.J., 2014. *An introduction to the event-related potential technique*. MIT press.

Myrberg, S., 2010. *The intonational phonology of Stockholm Swedish* (Doctoral dissertation, Acta Universitatis Stockholmiensis).

Myrberg, S. and Riad, T., 2015. The prosodic hierarchy of Swedish. *Nordic Journal of Linguistics*, *38*(2), pp.115-147.

Nakamura, K., Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Nagumo, S., Kubota, K., Fukuda, H., Ito, K. and Kojima, S., 2001. Neural substrates for recognition of familiar voices: a PET study. *Neuropsychologia*, *39*(10), pp.1047-1054.

Näätänen, R., Gaillard, A.W. and Mäntysalo, S., 1978. Early selective-attention effect on evoked potential reinterpreted. *Acta psychologica*, *42*(4), pp.313-329.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R.J., Luuk, A. and Allik, J., 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, *385*(6615), pp.432-434.

Näätänen, R., Paavilainen, P., Rinne, T. and Alho, K., 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical neurophysiology*, *118*(12), pp.2544-2590.

Näätänen, R. and Kreegipuu, K., 2011. The Mismatch Negativity (MMN). *The Oxford Handbook of Event-Related Potential Components*, pp.143-157.

Novén, M., Schremm, A., Nilsson, M., Horne, M. and Roll, M., 2019. Cortical thickness of Broca's area and right homologue is related to grammar learning aptitude and pitch discrimination proficiency. *Brain and Language*, *188*, pp.42-47.

Novén, M., Schremm, A., Horne, M. and Roll, M., 2021. Cortical thickness and surface area of left anterior temporal areas affects processing of phonological cues to morphosyntax. *Brain Research*, *1750*, p.147150.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), pp.97-113.

Oostenveld, R., Fries, P., Maris, E. and Schoffelen, J.M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience*, *2011*, pp.1-9.

Peirce, J.W., 2007. PsychoPy—psychophysics software in Python. *Journal of neuroscience methods*, *162*(1-2), pp.8-13.

Peirce, J.W., 2009. Generating stimuli for neuroscience using PsychoPy. *Frontiers in neuroinformatics*, p.10.

Poeppel, D., 2003. The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech communication, 41*(1), pp.245-255.

Politzer-Ahles, S., Schluter, K., Wu, K. and Almeida, D., 2016. Asymmetries in the perception of Mandarin tones: Evidence from mismatch negativity. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(10), p.1547.

Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., Alho, K., Martinkauppi, S., Ilmoniemi, R.J. and Näätänen, R., 2001. Memory traces for words as revealed by the mismatch negativity. *Neuroimage*, *14*(3), pp.607-616.

Pulvermüller, F. and Shtyrov, Y., 2006. Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. *Progress in neurobiology*, *79*(1), pp.49-71.

Riad, T., 1998. The origin of Scandinavian tone accents. *Diachronica*, *15*(1), pp.63-98.

Riad, T., 2006. Scandinavian accent typology. *STUF--Sprachtypologie und Universalienforschung*, *59*(1), pp.36-55.

Riad, T., 2014. *The phonology of Swedish*. Oxford University Press.

Roll, M., Horne, M. and Lindgren, M., 2010. Word accents and morphology—ERPs of Swedish word processing. *Brain research*, *1330*, pp.114-123.

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

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

Roll, M., Söderström, P., Frid, J., Mannfolk, P. and Horne, M., 2017. Forehearing words: Pre-activation of word endings at word onset. *Neuroscience Letters*, *658*, pp.57-61.

Roll, M., 2022. The predictive function of Swedish word accents. *Frontiers in Psychology*, *13*.

Schremm, A., Novén, M., Horne, M., Söderström, P., van Westen, D. and Roll, M., 2018. Cortical thickness of planum temporale and pars opercularis in native language tone processing. *Brain and Language*, *176*, pp.42-47.

Söderström, P., Horne, M., Frid, J. and Roll, M., 2016. Pre-activation negativity (PrAN) in brain potentials to unfolding words. *Frontiers in Human Neuroscience*, *10*, p.512.

Söderström, P., Horne, M., Mannfolk, P., van Westen, D. and Roll, M., 2017. Tone-grammar association within words: Concurrent ERP and fMRI show rapid neural pre-activation and involvement of left inferior frontal gyrus in pseudoword processing. *Brain and Language*, *174*, pp.119-126.

Squires, N.K., Squires, K.C. and Hillyard, S.A., 1975. Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and clinical neurophysiology*, *38*(4), pp.387-401.

Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D. and Leahy, R.M., 2011. Brainstorm: a user-friendly application for MEG/EEG analysis. *Computational intelligence and neuroscience*, *2011*, pp.1-13.

Wakusawa, K., Sugiura, M., Sassa, Y., Jeong, H., Horie, K., Sato, S., Yokoyama, H., Tsuchiya, S., Inuma, K. and Kawashima, R., 2007. Comprehension of implicit meanings in social situations involving irony: A functional MRI study. *NeuroImage*, *37*(4), pp.1417-1426.

Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T. and Ackermann, H., 2005. Identification of emotional intonation evaluated by fMRI. *Neuroimage*, *24*(4), pp.1233-1241.

Woldorff, M.G., Hackley, S.A. and Hillyard, S.A., 1991. The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones. *Psychophysiology*, *28*(1), pp.30-42.

Xu, Y., Gandour, J., Talavage, T., Wong, D., Dzemidzic, M., Tong, Y., Li, X. and Lowe, M., 2006. Activation of the left planum temporale in pitch processing is shaped by language experience. *Human brain mapping*, *27*(2), pp.173-183.

Zachau, S., Rinker, T., Körner, B., Kohls, G., Maas, V., Hennighausen, K. and Schecker, M., 2005. Extracting rules: early and late mismatch negativity to tone patterns. *Neuroreport*, *16*(18), pp.2015-2019.

Zatorre, R.J. and Gandour, J.T., 2008. Neural specializations for speech and pitch: moving beyond the dichotomies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), pp.1087-1104.

Zora, H., Riad, T., Schwarz, I.C. and Heldner, M., 2016. Lexical specification of prosodic information in Swedish: Evidence from mismatch negativity. *Frontiers in Neuroscience*, *10*, p.533.

Appendix

Participant information

Gender	Age	First Language(s)	Place (region) grow up	Laterality Quotient
F	23	Swedish, German	Skåne	100
M	19	Swedish	Skåne	93.1
M	24	Swedish	Skåne	100
M	25	Swedish	Västra Götaland, Skåne	92.86
F	21	Swedish	Skåne	71.43
M	27	Swedish	Skåne	86.21
M	21	Swedish	Stockholm, Skåne	58.33
F	28	Swedish	Skåne, Copenhagen, Cyprus	100
M	19	Swedish	Skåne	100
M	37	Swedish	Skåne, Blekinge	76.92
M	21	Swedish	Skåne	85.71
M	23	Swedish	Skåne	100
M	27	Swedish	Skåne	85.19
F	22	Swedish	Skåne	92.59
M	25	Swedish, Thai	Skåne	83.33
F	22	Swedish	Skåne	100
M	25	Swedish	Skåne	100
F	24	Swedish, Danish	Skåne, Copenhagen	93.33
F	23	Swedish	Skåne	100
Mean	24			90.47
SD	4.04			11.67