Holistic and combinatorial processing of Swedish tone accents in the brain: An MMN study

Renata Kochančikaitė

Supervisor: Mikael Roll

Abstract

All words in Swedish have a particular melody. There are two tonal patterns that are observed in different prosodic and morphological contexts – accent 1 and accent 2. The tone of accent 1 is the default melody in Swedish and is said to be of intonational nature. The melody of accent 2, on the other hand, only occurs in presence of certain suffixes in the word, hence, accent 2 is claimed to be a property of suffixes. This thesis investigated how these accents are processed in the native speakers' brain. A hypothesis was raised suggesting that accent 1 is post-lexical and accent 2 is morphological. Two test implications were derived: 1) the melody of accent 1 is processed together with the word itself and in the brain they have a common, holistic representation, a so-called whole-word memory trace, and 2) accent 2 words are a feature of the suffix rather than of the entire word, so combinatorial processing is said to be involved when listeners hear accent 2 words.

These test implications were investigated by conducting an electroencephalographic experiment designed to elicit a type of brain response that is called mismatch negativity. To elicit a mismatch response, a series of identical words was played to the study participants but every now and then a different, deviant word was inserted into the series. Four such pairs of standard and deviant words were used in the experiment, each pair probing into the neurophysiological responses to a particular combination of accent and suffix. The results corroborated the hypothesis and an additional discovery was made – there might be a difference not only in the way the tone accents are stored in the memory, but also in the way the suffixes are integrated into word representations in the memory, depending on the type of suffix. Accent-neutral suffix, also analysed as a clitic morpheme, elicited responses that point towards whole-word storage, whereas accent 2-inducing suffix elicited responses indicative of combinatorial processing.

Keywords: tone accents, Swedish, phonology, morphology, accent 1, accent 2, MMN, mismatch negativity

Acknowledgements

Table of Contents

List of Figures

List of Tables

1. Introduction

Spoken language and melody are virtually inseparable. Swedish represents one of those languages where different kinds of words have their own melodies – tone accents. Like all learning from experience, if some specific melody frequently co-occurs with a particular word type or word part, that melody is probably going to etch into the listeners' memory and contribute to distinct processing patterns in their brains. The current thesis is an attempt to shed more light on those patterns.

Swedish language has two tone accents – accent 1 and accent 2 – which occur in different morphophonological contexts. Even though their distribution looks complicated, clear regularities exist in assignment of tone accents. Some regularities pertain to prosodic facts while others seem to concern morphology and the mental lexicon. This distinction may have corresponding consequences for language processing in the native speakers' brains. The aim of the current thesis is to test a dual hypothesis which suggests that accent 1 is processed in whole-word form together with the stem it co-occurs with, whereas accent 2 is processed as a feature of a morphological unit and is decomposed in a sequential manner. These suggestions are in line with evidence from previous neurolinguistic research (Söderström, Roll, & Horne, 2012; Roll, Söderström, & Horne, 2013; Söderström, Horne, Mannfolk, et al., 2017; Roll et al., 2015; Söderström, Horne, Frid, & Roll, 2016; Schremm et al., 2018) as well as phonetic experiments (Engstrand, 1995, 1997) and phonological theory on Scandinavian accents (Myrberg & Riad, 2015; Riad, 2006, 2009, 2012).

1.1. Swedish tonal accents

Swedish language has a tone accent system where all words have a specific melodic pattern, associated with the primary stressed syllable. Two types of such melodies, i.e. two tone accents, are recognised in Swedish: accent 1 and accent 2. These accents have different surface realisations and occur in different morphological contexts. Sometimes tone accent is the only difference helping to distinguish meanings of words that constitute minimal pairs (according to Elert (1972), there are about 350 such minimal pairs in Swedish).

The pitch contours of accent 1 and accent 2 vary across dialects. In Central Swedish, which is the dialect investigated in the current thesis[1], the surface realisation of accent 1 is characterised by low-pitched tone on the stressed syllable (L*) and accent 2, by high-pitched tone on the stressed syllable (H*L) (Riad, 2006). In addition, these main tonal features are modified by the level of prominence with which the tone-bearing word is pronounced (Bruce, 1977). The L* (accent 1) and H*L (accent 2) tonal contours apply for words that occur in non-focal position, i.e. when they are pronounced with little prominence in the sentence. In focal position, i.e. when pronounced with greater prominence, an upwards movement of pitch (also called focal accent, or sentence accent) is added after the stressed syllable and the accents' surface realisation becomes L*H for accent 1 and H*LH for accent 2 (Riad, 2006).

There is no simple rule that defines which of the two tone accents a word will have. Accent assignment involves many aspects, including morphology and prosody. For example, monosyllabic words have accent 1 (Bruce, 1977: 18; Lahiri, Wetterlin, & Jönsson-Steiner: 62, 2005; Riad, 2009: 3), a regularity that pertains to a prosodic feature of the word. Another example where prosody plays a role is words with multiple stressed syllables, such as compounds – a majority of those have accent 2 in Central Swedish (Gårding & Lindblad, 1973: 41; Bruce, 1977: 19; Riad, 2006: 51). Morphology comes into play when the root of the word is affixed. Some prefixes force roots that otherwise bear accent 2 to change their tonal structure to accent 1, e.g. $_{acc2}$'tal-a 'to speak' but $_{acc1}$be-'tal-a 'to pay'. A lot of suffixes[2], on the other hand, impose accent 2 even on accent 1 roots, e.g. $_{acc1}$glad 'happy' but $_{acc2}$glad-are 'happier', $_{acc1}$bil 'car' but $_{acc2}$bil-ar 'cars'. In more complex words, multiple accent-assigning features may have to compete (Riad, 2009), e.g. $_{acc2}$be-'tal-ka-ˌnal 'payment channel' begins with the accent 1-inducing prefix be- but ultimately has accent 2 because its prosodic structure – marked by presence of two stresses – overrules the accent 1-inducing affix.

The examples listed above represent only a tiny portion of the Swedish tone accent landscape. There are more prosodic and morphological conditions that govern the distribution of tone accents in Swedish as well as more than one way to analyse the phonological data; the Background chapter features a more detailed review. However, one prominent pattern

---

[1] Central Swedish is the object of this study because it corresponds roughly to standard Swedish (Riad, 2012: 1353) and is the most well-studied Swedish dialect (Myrberg & Riad, 2015: 117) providing a sound body of research to rely on.

[2] Except the definite article suffix which cliticises to word roots without affecting their tonal structure (Riad, 1998: 65). Riad treats postposed definite article as clitic. In the current thesis, however, it is called a suffix for the sake of simplicity.

concerning compound words deserves some immediate attention, since it leads towards the hypothesis that motivates the current thesis.

1.2. Motivation, hypothesis, and research questions

As mentioned, due to secondary stress, compounds can only have accent 2 in Central Swedish. Consequently, noun stems that bear accent 1 have a limited set of continuations, i.e. a handful of suffixes that combine with accent 1 stems in a felicitous way. Nouns that begin with a stem that bears accent 2, on the other hand, can continue not only with certain suffixes but other root morphemes as well. Compounding is a productive process in Swedish that employs many word classes including nouns, proper nouns, adjectives, verbs, pronouns, numerals etc. (Teleman et al., 1999). Thus, the amount of possible continuations for accent 2 stems is much larger than that of possible continuations for accent 1 stems. In fact, Söderström, Horne, Frid, & Roll (2016) have used a lexicon database and estimated that accent 2 stems have about 11 times more continuations, most of them being compounds.

This unevenness in the morphological distribution of tone accents appears to have an impact on word processing in the brain, as revealed by a number of electroencephalographic (EEG) and neuroimaging experiments. It has been shown that valid words with suffixes associated with accent 1 yield faster reaction times in semantic tasks than valid words with suffixes associated with accent 2 (Söderström, Roll, & Horne, 2012). Also, listeners, when asked to guess how a cut-off word continues, were faster and more accurate when presented with noun stems with accent 1 rather than accent 2 (Söderström, Horne, & Roll, 2017). This holds even if the semantic content is absent and the word frequency condition is equal – Söderström, Horne, Mannfolk, et al. (2017) and Söderström, Horne, & Roll (2017) used made-up words (hence $f_{word}=0$) that were felicitous combinations of fake stems and real suffixes instead of actual Swedish words. Their findings support the claim that accent 1 has superior predictive capacity relative to accent 2. As suggested already by Roll et al. (2015) and Söderström, Horne, Frid, & Roll (2016), accent 1 stems, having only a small set of valid continuations, may more efficiently pre-activate the memory traces of felicitous suffixes in the brain and successful pre-activation should speed up the processing of the entire word. Schremm et al., (2018) also provided evidence that tones may be processed as part of stored whole-word representations for familiar speech segments, such as inflected word forms. Whole-word storage should indeed facilitate rapid processing, however, it is not clear whether it applies to accent 1 word forms and accent 2 word forms alike. Accent 1 has few legitimate suffix

continuations, so accent 1 word forms are highly predictable and therefore likely to have developed their own whole-word memory traces. Accent 2, in the meantime, seems to be processed differently. The suggestion raised in this thesis is that it could be stored in the mental lexicon as a part of accent 2-inducing suffixes. In that case, accent 2 word forms would be treated as morphosyntactic sequences and processed in a rule-based combinatorial manner.

The idea of lexical storage of accent 2 in suffix morphemes comes from phonological literature (Riad, 2012). According to Riad (2006, 2009) and Myrberg & Riad (2015), there is an essential difference between accent 1 and accent 2, namely, that accent 2 has a lexical tone, which accent 1 lacks. The pitch contour of accent 1 words can be fully explained with prosodic factors such as word prominence tones or phrase boundary tones (Riad, 2006). But, when the melodic contributions arising from prosodic circumstances are stripped off of simplex accent 2 words, there still remains a high tone (H*) in the surface realisation of those words (Riad, 2006). That high tone is treated as intrinsic to the morphemes that assign accent 2 to a simplex word; this is called lexical assignment. Prosodically defined accent melody is not intrinsic to any morpheme or lexeme, so accent 1 is not considered a lexical tone. Instead, it is said to be assigned post-lexically, i.e. after the morphemes and lexemes have been retrieved from the mental lexicon and stringed together into a maximal prosodic word (Myrberg & Riad, 2015). As can be seen, this theoretical account is compatible with the accumulating experimental evidence that accent 1 and accent 2 are processed differently. The opposition of intonational nature of accent 1 and morphosyntactic nature of accent 2 is thus in line with whole-word retrieval of accent 1 word forms versus decompositional processing of accent 2 word forms.

Lexical retrieval and morphological decomposition as two alternative ways of processing of related word forms has been studied before. It was postulated, among others, by Pinker (1991) who advocated a dual manner of processing of regular and irregular past tense verbs in English and called it the dual-route model. According to this model, irregular verb forms are memorised in their entirety, while regular verb forms are composed online by following a morphological rule. Pinker's model has received corroboration from multiple EEG experiments done on multiple languages (e.g. Morris & Holcomb, 2005; Newman, Ullman, Pancheva, Waligura, & Neville, 2007; Bakker, MacGregor, Pulvermüller, & Shtyrov, 2013), including Swedish (Schremm, Novén, Horne, & Roll, 2019). The dual-route model has since been developed further: for example, a reaction-time study of Dutch nouns indicated holistic lexical storage even for completely regular noun forms, but only those that yielded faster reaction times (Baayen, Dijkstra, & Schreuder, 1997). In this case, the researchers attributed

the differences in processing speed to surface frequency. When it comes to Swedish, however, accent 1 words yield faster reaction times even if the surface frequency is the same. Consequently, the difference in processing speed of Swedish words might indicate that different processing routes are invoked depending on the type of tone accent. The likelihood that different tone accents initiate different processing routes is also compatible with Riad's (2009: 207) position that accent 2 is a property of suffixes while accent 1 is superimposed post-lexically.

To summarise, suggestions from Riad's phonological theory and from experimental evidence on Swedish tone accents seem to converge into an idea of non-equivalent processing of accent 1 and accent 2 that goes along the lines of the dual-route model of word processing: lexical, holistic storage of accent 1 words and rule-based decomposition of accent 2 words. Hence, a twofold **hypothesis** is formulated in the current thesis, suggesting that:

- accent 1 is associated post-lexically with stems and is retrieved from the lexical memory along with the stem, in whole-word form;

- accent 2 is associated lexically with suffix morphemes and is combined with the word stem when a combinatorial rule is followed to attach the suffix to the stem.

To the best of the author's knowledge, such claims have not been tested before. The purpose of this thesis is thus to conduct a suitable experiment to test them and thereby to contribute to the empirical investigations of the cognitive nature of tone accents in Central Swedish. This work is guided by the following **research questions**:

1. Does the MMN paradigm support the previous experimental evidence about the likelihood that accent 1 is represented holistically in whole-word forms?

2. Is it likely that accent 2 is represented in the brain as a feature of suffix morphemes, as proposed by Riad's phonological model?

1.3. Method

The methodological side of the current investigation relies on a neurophysiological pattern of brain response that is called mismatch negativity (MMN). An MMN response is observed in auditory oddball experiments, i.e. when a series of repetitive standard stimuli is disturbed by a deviant stimulus that occurs randomly at a low probability. In such experiments, deviant stimuli elicit a negative peak of the EEG waveform at 100-250 ms after onset of the stimulus

(Näätänen, 2001). This method is suitable for testing the current hypothesis because the elicited MMN amplitude also depends on different linguistic, including lexical and syntactic, properties of a deviant stimulus (Pulvermüller & Shtyrov, 2006). If the deviant stimulus is an actual word, known to the participant, it will elicit greater MMN amplitude than a deviant stimulus that is a pseudoword, not known to the participant (Korpilahti, Krause, Holopainen, & Lang, 2001; Pulvermüller, 2001). This type of response is called *lexical MMN* and has been associated with automatic activation of a lexical memory trace of a known word (Pulvermüller & Shtyrov, 2003). According to the current hypothesis, accent 1 word forms should activate whole-word memory traces and elicit lexically enhanced MMN. An opposite modulation of the negativity amplitude of the brain response is called *Syntactic MMN* and has been associated with syntactic priming. If the deviant stimulus is a well-formed, grammatical sequence, it will elicit a lesser MMN peak compared to a deviant stimulus that violates some compositional rule (Pulvermüller & Shtyrov, 2003; Pulvermüller & Assadollahi, 2007; Hanna et al., 2014). The explanation behind syntactic MMN is that elements, represented by separate memory traces that can form valid grammatical sequences, will prime each other – and priming reduces the absolute size of event-related potentials (Holcomb & Neville, 1990). Accent 2 words then, as they are hypothesised to be decomposed into morphemes, should activate the memory traces of those morphemes and elicit syntactically reduced MMN.

Even though the majority of research with lexical and syntactic MMN seems to be carried out on tonally unspecified data, lexical tone languages have also been successfully investigated with this method. Lexical MMN responses were observed by Yue et al. (2014) who tested memorisation of novel segment-tone patterns as single lexemes in Mandarin Chinese speakers. Syntactic MMN was elicited by Chang, Lin, & Kuo (2019) who compared Mandarin Chinese and Taiwanese native speakers' responses to conditions where tone substitution was required by tone sandhi rules. Phonological aspects of pitch accent languages are also not virgin in this methodological area: Zora, Riad, Schwarz, & Heldner (2016) and Zora, Riad, & Ylinen (2019) have tested interaction of stress and morphology in Swedish and found both lexical and syntactic MMN. However, the current study appears to be the first to subject Swedish tone accents to an MMN experiment to investigate the differences in their processing in the brain.

1.4.Structure of the thesis

The thesis is structured in six main chapters, Introduction being the first. It is followed by the Background chapter where two main theoretical models of Swedish tone accents are presented. A discussion on the complementary and conflicting aspects of these models is given within the scope that is relevant for the current study. Third is the Method chapter in which the experimental design is described and information about the subjects and experimental procedure is given. This chapter also features a description of data processing pipeline and motivation for the choice of statistical methods for data analysis. Fourth is the Results chapter, it presents the outcome of the visual and statistical analysis. It is followed by Discussion and Summary chapters. At the end of the thesis, an Appendix with complementary statistical data is attached.

2. Background

This chapter consists of two main parts. Section 2.1. presents two classical theoretical accounts of Swedish tone accents – the Lund model and a model advocated by Tomas Riad. The highlights of this discussion are: lexical and phonological nature of the tone accents, equipollent and asymmetric treatment of the tone accent opposition, and approaches to the accent system from the mental lexicon point of view and surface representation point of view. Section 2.2. is dedicated to the theoretical aspects of the method of the current study. Subsection 2.2.1. gives a short overview of the electroencephalography (EEG) and informs about important technical aspects and necessary theoretical assumptions that need to be kept in mind when conducting an EEG experiment. In 2.2.2., the event-related potential (ERP) technique is presented and a selection of relevant ERPs is briefly described. Subsection 2.2.3 is a review of the mismatch negativity paradigm and its applications in neurolinguistic research.

2.1. Prosodic accounts of the Swedish tone accent system

Two analyses stand out in the phonological literature on Swedish tone accent system – the Lund model (Bruce & Gårding, 1981; Bruce, 1977; Heldner, 2001; Ambrazaitis, 2009) and what could be called Riad's model (Riad 1998, 2006, 2009, 2012; Engstrand 1995, 1997). The Lund model describes the accent dichotomy in sentence context, looking at how tone accents are realised in speech and how they interact with intonation phenomena. Riad's model, on the other hand, puts the grammatical and diachronic context in focus and examines what morphophonological elements induce one or the other tone accent on a given word. Proponents of both models seem to agree about many of the prosodic regularities observed in accent-bearing words – for example, that different levels of intonational prominence affect the tonal pattern of both accents equivalently (Bruce, 1977; Heldner, 2001; Myrberg & Riad, 2015) or that the domain of tone accents is not lexical words but prosodic words (Bruce, 1977; Ambrazaitis, 2009; Myrberg & Riad, 2015). However, there is an essential difference between the two models; it concerns the view on the (non)lexical nature of the tone accents and what the pure tonal patterns of both accents look like when isolated from intonational factors.

In the Lund model, accent 1 and accent 2 are both considered lexical and share the same tonal representation (Bruce, 1977; Gussenhoven & Bruce, 1999: 237). The pure accent representation reveals itself in non-focal positions, and according to the Lund model, in Central Swedish it is a fall from high tone to low tone (HL). The difference between the two accents

lies in their relative timing with regard to the primary stressed syllable of the word – the fall from H to L starts earlier on accent 1 words than on accent 2 words. An important result of this delay is that different tones end up on the stressed syllable. In accent 1 words, the fall takes place before the stressed syllable (even before the start of the word, if it is stress-initial (Bruce, 1977)), so the stressed syllable aligns with the low tone (HL*). In accent 2 words, on the other hand, the fall takes place during the stressed syllable, so the stressed syllable aligns with the high tone (H*L). The same timing delay applies to words in focal position where a final rise to the so-called focal H tone is attached yielding a HL*H tone pattern for accent 1 words and a H*LH tone pattern for accent 2 words. The Lund view can therefore be summarised as equipollent because it assumes, timing aside, equivalent surface representations for both accent 1 and accent 2 (Ambrazaitis, 2009: 47).

In Riad's model, the pair of tone accents is treated as privative, i.e. asymmetric in terms of markedness: accent 2 is the marked member of the pair since its tonal pattern has a lexical origin[3], whereas accent 1 is unmarked because its tonal pattern is explained solely with prosodic (=post-lexical) phenomena (Riad, 2006). In focal position, the tonal pattern for accent 1 words is L*H and H*LH for accent 2 words; the shared portion of the tonal pattern – the pitch rise LH – is identified as phrasal prominence tone. Consequently, according to Riad's model, only accent 2 has a pure tonal representation, a high tone associated with the stressed syllable (H*), while accent 1 is left with Ø and is realised in non-focal position as a generic intonational L*.

In a way, 'privative' does not necessarily oppose 'equipollent'. Bruce (1998) and Gussenhoven & Bruce (1999: 244) admit that accent 2 is the marked member of the pair and accent 1 is default, as is seen from the fact that unintegrated loanwords receive the tonal pattern of accent 1. Ambrazaitis (2009: 47-48) points out that both models could be understood as intersecting in a non-contradictory way because their assumptions operate on different planes – the Lund model is primarily concerned with phonetic realisation and phonological representation of the tone accents, while Riad's model is centred on lexical specification of the accents in relation to how they are perceived and processed. Yet, Riad (2009: 206) opines that

---

[3] It must be noted, though, that instantiations of accent 2 in words with multiple stresses, typically compounds, are considered post-lexical in this model because in such words the HL*H pattern is a result of a prosodic circumstance (Riad, 2006). It is only in the non-compound words where the lexical component of accent 2 manifests itself – and in those cases it is usually assigned by suffixes (Riad, 2012).

markedness should preferably be reflected in representation[4] and thus disagrees with this reconciliation plan.

In Riad's phonological account, the link between the phonetic facts and markedness is accent 2 being 'a property of suffixes' (Riad, 2009: 207). He explains in a simple manner: at Step 1, all Swedish words with stress on the final syllable (including monosyllables) can be put aside from the discussion because it is commonly agreed that they do not get accent 2 because of prosodic limitations (e.g. Ambrazaitis, 2009; Bruce, 1977; Myrberg, 2010; Riad, 1998); at Step 2, the remaining words can be sorted into two groups – those that get accent 1 due to prosodic anacrusis (existence of an unstressed syllable before the stressed one in the word) and those that get accent 2 due to the presence of an accent 2-inducing suffix. Riad (2009) discusses various morphophonological combinations where anacrusis and accent-inducing affixes are competing within the same word and shows systematically that accent variation follows the suffix morpheme.

Contrary to that, Bruce (1977: 17) argued that the main factor governing tone accent distribution in non-compound words is the placement of stress, and subsumed morphological composition, including suffix types, under the prosodic scheme of accent prediction. However, in support of Riad's account, it can be pointed out that the set of suffixes that assign accent 2 is stable across dialects despite dialectal variation of the surface representation of the Swedish tone accents (Riad, 2009: 210). Accent 1, on the other hand, is subject to more variation pertaining to intonational specificity of a given dialect. This points at the morphological nature of accent 2, where its lexical status is said to originate. Furthermore, some Swedish dialects have lost the lexical tone distinction by reducing accent 2 to the same tones that occur on accent 1 words (Riad, 2006: 42). This goes in line with the claim that the tonal pattern of accent 1 is the default one and is intonational in its nature.

Riad's model has gained some support from phonetic studies by Engstrand (1995, 1997). Engstrand (1995) recorded how native speakers of Central Swedish read the words acc1*länderna* 'the countries' and acc2*länderna* 'the loins' in focal and non-focal positions, both in sentence context and as single words. Then he measured the pitch changes in the target words across all conditions and found out, among other things, that a) the tone fall during the accent 1 word was significantly smaller, shorter and less uniform than the fall during the accent 2 word,

---

[4] For the current study, it is also advantageous if the theoretical background allows a more direct path of connecting empirical observations on language processing to abstract categories employed by the theoretical models.

and b) that the tone fall during the accent 1 word in non-focal position was only present when preceded by a high tone on the previous (focused) word in the sentence. This led him to conclude that whether the HL pattern would manifest for accent 1 words was predictable on prosodic grounds, whereas HL pattern for accent 2 was stable despite variations in intonational context. In the companion study, Engstrand (1997) recorded spontaneous speech and compared the pitch contour of the recorded accent 1 and accent 2 words. The results were in line with the controlled speech experiment, as accent 2 words consistently showed a tone fall on the stressed syllable, while the tone patterns on the accent 1 words were quite variable.

2.2. Understanding the method

The current study aims to provide neurophysiological evidence on the nature of Swedish tone accents. The chosen method is electroencephalography which, despite becoming more and more widely used for linguistic research, perhaps is not quite yet part of the general linguistic profile. Therefore, a concise introduction to EEG is included here, to serve as a firm ground on which informed methodological decisions could be made and appropriate interpretation of the experimental data could be ensured.

2.2.1. Electroencephalography

Electroencephalography (EEG) is a technique that allows to measure electrical brain activity. It is often used in neurolinguistic research, thanks to its non-invasiveness, great temporal resolution, and low costs relative to other neuroimaging techniques. During an EEG session, a set of electrodes is placed on the scalp in designated positions and good connectivity between them and the skin is ensured. The subject is then exposed to a series of stimuli, while the neurophysiological responses to these stimuli are registered as voltage potentials at the locations of the electrodes, amplified, and digitally recorded. EEG recordings can have a temporal resolution of hundreds or thousands of data points per second. The electrodes are very sensitive and, with the help of an amplifier, minute voltage changes can be distinguished. Together this allows almost immediate registration of subtle changes in neurophysiological response – a great advantage of EEG. An EEG recording consists of an array of brainwaves, i.e. voltages plotted over time for each electrode, along with time-stamps for every moment when a stimulus was presented to the study subject (also called stimulus onset time).

A brainwave can be understood as a summation of underlying components, where each component contributes to the overall waveform, together shaping its amplitude peaks and valleys. Once the brain's responses to the same stimulus are averaged, the background noise is mitigated and the responses emerge that have been elicited by that stimulus. These responses are called event-related potentials (ERP) and are particularly relevant for neurolinguistic research. ERP components can be described as patterns of voltage potential changes that are functionally related to an experimental variable or a combination of those (Donchin, Ritter, & McCallum, 1978: 353). ERP components are presumed to represent sensory, cognitive, affective, and motor processes generated by distinct neuronal populations in the brain. Components caused by muscle activity, breathing, sweating, or external electric fields are considered noise artefacts, as they are usually not related to presentation of the experimental stimuli or mental processes triggered by them. Therefore, ERP data must be carefully filtered and artefacts must be compensated for before any functional analysis can be performed. Unfortunately, artefact removal results in loss of data for trials with contaminations that are beyond repair. That is one of the reasons why many trials and repetitions of the same stimulus are needed in the experimental design. Another reason is that a lot of the content in the EEG signal is background neural activity, unrelated to the cognitive processes that the experimenter is trying to probe into. To enhance the signal from ERPs related to experimental variables, the recorded EEG responses need to be averaged before analysis. Averaging is considered a sufficient means to achieve this because the background activity is presumably random (Kaan, 2007: 572), so the fluctuations are taken to simply cancel each other out.

Spatial resolution of EEG recordings is much more limited than temporal resolution. For one, the number of spatial data points depends directly on the amount of electrodes used, which typically is 32, 64, or 128. But, more importantly, EEG data does not reveal where exactly in the brain the event-related potentials have originated and which populations of nerve cells have caused them. EEG signal is a summation of lots of separate post-synaptic and action potentials originating at neural membranes (Nunez, Nunez, & Srinivasan, 2019: 196). Modelling of intracranial sources of EEG signal on this microlevel is complicated, so a simplified model based on the concept of equivalent current dipoles is usually employed. A dipole represents the summed activity of neurons that are near each other (Luck, 2014: 43), which is one way of approximating cortical current sources in superficial (up to 0.5 cm inwards) layers of the cortex (M. D. Nunez & Srinivasan, 2016: 176). The current study, however, did not have any specific hypotheses about sources, so dipole reconstruction was not carried out.

Despite having poor spatial resolution, EEG provides a means for investigating spatial distribution of the brain responses. A topographical scalp map showing the average polarity and amplitude of potentials recorded at the electrode locations at a selected time point can be obtained. Even though it does not reveal the actual location of the cortical generators of the potentials, scalp distribution is an important defining attribute of ERP components (Donchin et al., 1978; Luck & Kappenman, 2012; Spencer, Dien, & Donchin, 2001). Apart from scalp distribution, many scholars consider the morphology (i.e. the temporal and amplitude dynamics of ERP waveforms) an important feature of an ERP component (e.g. Donchin & Isreal, 1980; Kiesel, Miller, Jolicœur, & Brisson, 2008; Spencer et al., 2001).

From the theoretical point of view, the combination of timing, amplitude polarity, scalp distribution and sensitivity to manipulations of experimental variables is not a sufficient formal definition of an ERP component (Kappenman & Luck, 2012). Amplitude peaks and latencies of the same ERP components vary with slight alterations of experimental conditions (Kiesel et al., 2008), while scalp distributions are affected, among other things, by the choice of reference electrode (Kappenman & Luck, 2012). Interpretation of ERP data is further complicated by the great likelihood that multiple ERP components overlap in the same EEG wave and smear out the visible morphological and topographical features (Kappenman & Luck, 2012). It is actually difficult to ever be sure whether one really has detected the components that, considering their morphology and scalp distribution, seem to match the expectations of the experimental setting. However, Kappenman and Luck (2011) maintain that uncertainty is a companion of many sciences, and propose the 'converging evidence approach' as the best practical way of interpreting ERP data. In the converging evidence approach, all the above-mentioned factors – timing, polarity, amplitude, scalp distribution, experimental conditions – are taken into consideration and the strength of evidence for the probable ERP components is evaluated. This approach is also adopted in the current work.

2.2.2. *Event-related potentials*

With the help of dedicated digital tools, measurement of the physical characteristics of the recorded EEG waves is quite straightforward. Evaluating the possible interplay of the various ERP components, on the other hand, requires adequate knowledge and human reasoning to

apply that knowledge appropriately. A concise review of major ERP components the way they are observed in healthy humans[5] is therefore in order.

The nomenclature of ERP components is not consistent but, typically, the name of a component starts with a letter 'P' or 'N', indicating whether the waveform is positive-going or negative-going, and is followed by a number which either indicates the latency of the peak associated with that ERP component (e.g. P600 refers to a positive peak at around 600 ms after stimulus onset) or the ordinal number of the peak (e.g. N1 refers to the first negative peak in the ERP wave after stimulus onset) (Luck 2014: 72-73). Some ERP components are named after the neural processes that they are presumed to reflect (such as *Novelty P3*, indicating adaptation of the mental model of a known stimulus to a novel stimulus (Debener, Kranczioch, Herrmann, & Engel, 2002)), while others hint at the type of experimental setup that is required to evoke them (e.g. *mismatch negativity*, which refers to the oddball paradigm where a mismatch between rare and common stimuli is perceivable (R. Näätänen, Gaillard, & Mäntysalo, 1978)).

Perhaps the earliest language-related ERP components is *early left anterior negativity* (ELAN). It is observed as early as 100-300 ms after stimulus onset time and, as the name suggests, is distributed over anterior electrodes, often lateralised to the left (Swaab et al., 2012: 427). Functionally, ELAN is elicited by word category violations (Rösler, Pütz, Friederici, & Hahne, 1993). A very similar but later component is *left anterior negativity* (LAN), which shows up around 300-500 ms after stimulus onset (Morris & Holcomb, 2005: 964). Kutas & King (1995) treat anterior negativity at both latencies as a single component that is linked more generally to working memory storage employed during verbal processing. But, according to a 3-stage model of linguistic processing, proposed by Friederici (1995) and Friederici & Weissenborn (2007), ELAN and LAN are separate components: ELAN represents integration difficulties during the early stage, when (contradictory) word category information is detected, whereas LAN indicates difficulties during the middle stage, i.e. when (incongruent) morphosyntactic information is processed and integrated. Friederici & Weissenborn (2007) note that the processes presumably underlying the ELAN/LAN complex are affected by the level of morphological complexity of the language under investigation – higher ELAN/LAN amplitudes are expected for languages with more complex morphology. Indeed, experiments

---

[5] A lot of EEG research is conducted in clinical circumstances, in relation to mental and neural disorders but ERP component modifications observed in special populations are outside the scope of the current work.

with Swedish, a language that employs morphology, suggest a subtype of early anterior negativity called *pre-activation negativity* (PrAN) which is related to prediction of specific morphological information (Söderström et al., 2016). PrAN has been observed over frontocentral electrode sites at 136-280 ms after stimulus onset when study subjects anticipated a specific suffix cued by a stem tone (Söderström et al., 2016).

In more general contexts one can encounter the names N2 or N200 – umbrella terms referring to a complex of ERP components that share the same polarity and latency (negative, peaking around 200-350 ms after stimulus onset) but differ in their scalp distribution, are related to many different cognitive processes, and are elicited by different experimental setups (Folstein & Van Petten, 2008). One of these components, elicited by a passive oddball paradigm in auditory modality, is the *mismatch negativity* (MMN), a frontocentral negativity already presented in the Introduction and further discussed in one of the later subchapters. Other components that belong in the N2 category concern non-auditory modalities and are highly unlikely to be elicited by the experiment in the current work and; for that reason, these components are not listed here.

Another language-related ERP component is N400 which reflects semantic processing and integration. N400 manifests as a large negative wave peaking at around 400 ms that shows a broad scalp distribution (M Kutas & Hillyard, 1980) and centroparietal maximum (Swaab et al., 2012: 400). It is usually elicited with complete phrases or sentences or even short narratives, but stand-alone words have also been shown to elicit N400. Also, real words have been shown to elicit smaller N400 amplitude than pseudowords, and random strings of letters do not elicit N400 at all (Swaab et al., 2012).

Probably the latest, in terms of latency, well-studied language-related component is P600, also called *syntactic positive shift* (SPS) (Hagoort, Brown, & Groothusen, 1993). This ERP component is associated with re-evaluation of information that is inconsistent with the syntactic structure expected by the subject (Osterhout & Holcomb, 1992) or with processing of congruent but highly complex or unusual syntactic sequences (Swaab et al., 2012: 419). This positivity is usually a clear and robust wave that starts between 300 and 500 ms after stimulus onset, lasts for several hundred milliseconds, and has a widespread scalp distribution and largest amplitudes over centroparietal areas (Morris & Holcomb, 2005: 964).

2.2.3. Mismatch negativity – a tool for linguistic inquiry

Of all ERP components, MMN occupies a central place in this thesis. Primarily, MMN is a reflection of memory operations in auditory perception, although visual and tactile stimuli have also been reported to elicit MMN responses (Kujala, Tervaniemi, & Schröger, 2007: 2). Elicitation of MMN is done in 'oddball' paradigm – an experiment protocol where a series of repetitive standard stimuli is randomly interrupted with rare (usually occurring with 15-20% frequency) deviant stimuli. Average responses to standard stimuli are subtracted from those to deviant stimuli to obtain an MMN curve where a negativity increase can be observed. During repeated exposure to standard stimuli, the hearer learns the regularities of the input and the memory trace of the acoustic shape of that input is active in the auditory system. Deviant stimuli do not match that memory trace as they violate the pattern or rule shared by the standards (Näätänen, 2001). When the auditory comparison process detects a mismatch, an increased negativity in frontocentral recording sites is evoked at 100-250 ms after stimulus onset (Näätänen, 2001). This comparison of memory traces and new sensory input takes place even without the subjects' active attention to stimuli (Näätänen, 2001), which allows to utilise MMN in passive experimental designs. Moreover, it has been shown that the subjects do not need to be aware of an acoustic difference between standards and deviants (Van Zuijen, Simoens, Paavilainen, Näätänen, & Tervaniemi, 2006); MMN has been recorded even in coma patients (Näätänen, 2001). This shows that the MMN response is, to a great extent, automatic[6].

Like many ERP components, MMN exhibits a certain amplitude, latency, and topography variation. MMN amplitude can be modulated by, e.g., active attention to the stimuli, variations in frequency and other acoustic features of the stimuli, the subjects' hearing (Näätänen, 2001), age (Strömmer, Tarkka, & Astikainen, 2014), medical conditions etc. These are the factors that a linguist usually wants to keep constant in the experimental design. But, in addition, MMN is sensitive to a variety of abstract linguistic features including phonological, lexical, semantic, and syntactic ones (Pulvermüller & Shtyrov, 2006), as well as word frequency-related aspects (Alexandrov, Boricheva, Pulvermüller, & Shtyrov, 2011). These factors are in a linguist's particular interest and can be varied in MMN experiments to shed

---

[6] However, timing is important in MMN experiments. The mismatch detection utilises memory traces generated by an earlier feature tracking process which operates at the sensory level (Näätänen & Winkler, 1999). These traces usually fade after 5-10 seconds after stimulation (Näätänen, 2001). Despite overlap with other ERP components, spacing the stimuli 1-2 seconds seems to be a reliable interval to maintain elicitation of MMN (Mäntysalo & Näätänen, 1987).

more light on the neurophysiological correlates of language processing. Therefore, a survey of various kinds of MMN responses in linguistic contexts is presented below.

On phoneme level, Koyama et al. (2000) demonstrated that known language sounds elicited larger MMN responses than sounds not associated with any familiar phoneme. Cheour et al. (1998) followed how infants acquire language and reported that 6-month-old Finnish infants showed larger MMN responses to deviants that had greater acoustic difference from standards. At the age of 1 year, however, the same infants showed larger MMN responses to deviants that crossed a phonemic boundary in Finnish, despite having lesser acoustic difference from standards. Bonte et al. (2005) demonstrated MMN sensitivity to language-specific phonotactic rules. Recently, Chang, Lin, & Kuo (2019) subjected phonological rules governing substitution of lexical tones (tone sandhi) to an MMN experiment and discovered a latency change in MMN responses to phonological differences between Chinese and Taiwanese lexical tones.

2.2.3.1.Lexical mismatch negativity (lMMN)

MMN can also reflect access to lexical memory. In experiments where one syllable is repeated as standard and an unrelated syllable comes up as deviant, smaller MMN is elicited, whereas if the deviant syllable completes the previous standard syllable so that together they form a real word, the MMN response is larger (Pulvermüller et al., 2001). This response pattern is also valid for stimuli that are complete words. Enhanced MMN response to lexical units, known as *lexical MMN*, has been elicited in multiple experiments (e.g., Korpilahti, Krause, Holopainen, & Lang, 2001; Kujala et al., 2002; Shtyrov & Pulvermüller, 2002) by deviants that were actual words known to the subject, compared to deviants that were meaningless pseudowords. In addition, Pulvermüller et al. (2001) have shown that lMMN is not caused by phonemic or acoustic properties of the word. Instead, it is agreed that MMN amplitude is enhanced by presence of a neural memory trace for known words, stored in whole-word form (e.g. Bakker et al., 2013; Hanna et al., 2017; Leminen, Leminen, Kujala, & Shtyrov, 2013; MacGregor & Shtyrov, 2013; Pulvermüller et al., 2001; Shtyrov & Pulvermüller, 2002). Such word memory traces are formed during the learning process as distributed, strongly interconnected circuits of neurons (Bakker et al., 2013; Shtyrov & Pulvermüller, 2002). Differently from sensory acoustic memory traces that decay within seconds (Mäntysalo & Näätänen, 1987; Näätänen, 2001), word memory traces remain in the long-term memory network and are robust enough to be activated in a passive oddball paradigm. When a known word is recognised and its mental

representation – word memory trace – is ignited, additional neural activity occurs and is registered as added negativity in the MMN response (Hanna et al., 2017: 88). Pseudowords, on the other hand, lack a pre-existing memory trace and the MMN response to them is smaller (Bakker et al., 2013: 188), i.e. the MMN amplitude of pseudowords is only conditioned by the acoustic differences between deviant and standard stimuli. On the other hand, lMMN for real words is modulated by word frequency. E.g., Alexandrov et al. (2011) performed a study on Russian language and compared MMN responses to a word with >200 instances per million words with a similar-sounding word with <1.5 instances per million words and corrected the results for acoustic differences. A significantly more pronounced MMN response was elicited by the frequent word compared to the infrequent one.

An interesting aspect suggested by research done on lMMN is that the contents of a whole-word memory trace can include more elements than just a meaningful string of phonemes, i.e. one lexical word. A study by Cappelle Bert, Shtyrov, & Pulvermüller (2010) has demonstrated lexical MMN enhancement for correctly formed discontinuous particle verbs in English, such as *heat up* and *cool down*, compared to infelicitous combinations as **fall up* and **rise down*. Hanna et al. (2017) have demonstrated the same for German particle verbs. These results not only complement the list of linguistic aspects that can be fruitfully investigated in MMN experiment protocol, but also bear theoretical implications, namely, that particle verbs may deserve to be treated as single lexical units (Hanna et al., 2017: 89).

Pitch information might also be part of a whole-word memory trace, as suggested by Yue, Bastiaanse, & Alter (2014). In their study, Yue and colleagues taught native speakers of Mandarin Chinese some novel segment-tone patterns that do not exist in Mandarin Chinese. MMN responses to these novel segment-tone patterns relative to real words were recorded before and after training. The results showed significantly enhanced MMN in post-training recording which led the authors to conclude that segmental information, i.e. phoneme sequence, and suprasegmental pitch information, i.e. lexical tone that co-occurs with that phoneme sequence, can be learned in integrative fashion and, consequently, stored in a single whole-word form as a long-term memory trace. These findings are important in the light of the current thesis because they go in line with previous relevant research that suggested that Swedish tone accents, too, may be stored along in whole-word forms (Schremm et al., 2018). Moreover, MMN paradigm has been successfully used to provide support for categorical perception of Chinese lexical tones (Xi, Zhang, Shu, Zhang, & Li, 2010) and integral processing of segmental and tonal information (Choi, Tong, Gu, Tong, & Wong, 2017). Considering the above, MMN is a promising tool to investigate whether integration of tonal

and segmental information on word form level holds for Swedish as well, as predicted by the first half of the hypothesis in the current work.

2.2.3.2. Syntactic mismatch negativity (sMMN)

The second half of the current hypothesis concerns combinatorial processing and syntactic sequences. It is known that MMN responses are sensitive not only to single-unit memory activations but also to interactions between words or morphemes that make up a sequence (Leminen et al., 2013). MMN responses in such contexts have been termed *syntactic MMN*. Its response pattern is opposite to that observed in lMMN conditions. While lMMN is larger for real words than pseudowords, well-formed morpheme or word combinations elicit smaller MMN compared to ill-formed sequences, as shown by a multitude of studies. For example, Pulvermüller & Shtyrov (2003) presented English subjects with the verb stem *come* and its inflected form *comes* in linguistic (preceded by pronoun *we*) and non-linguistic (preceded by noise) contexts. The listeners heard a grammatically correct syntactic sequence *we come*, an ungrammatical sequence **we comes*, and neutral *come* and *comes* in isolation. The authors found smaller MMN responses to the deviants that were syntactically correct relative to incorrect ones. Their conclusion was that sMMN reduction may be due to a priming effect (known to reduce the overall amplitudes of negative-going ERP components (Holcomb & Neville, 1990)); priming takes place in syntactic contexts, when morphemes that often occur together, pre-activate each other's memory traces.

For the German language, Menning et al. (2005) embedded three acoustically similar words *Riesen*, *Rasen*, and *Rosen* in the same sentence '*Die Frau düngt den RASEN/RIESEN/ROSEN im Mai*' (The woman fertilizes the lawn/giant/roses in May) so that the sentence with *Rasen* was fully grammatical, *Riesen* introduced a semantic error due to selectional restrictions of the verb, and *Rosen* – a morphosyntactic error due to violation of number agreement (*den* is a singular article while *Rosen* bears plural suffix). Both of the ungrammatical conditions elicited lexical and syntactic MMN responses which the authors interpreted as index of early access to syntactic and semantic information and fast detection of correctness, possibly through automatic categorisation processes.

It has been mentioned that lMMN is sensitive to lexical frequency; this goes in line with the long-term whole-word memory trace interpretation. sMMN, however, does not seem to be modulated by frequency of co-occurrence of specific morphosyntactic sequences, as demonstrated by Pulvermüller & Assadollahi (2007). In their study, ungrammatical word

sequences (*der Wut* and *die Mut*), elicited a significantly smaller MMN amplitude than grammatically correct word pair sequences (*die Wut* and *der Mut*), regardless of their probability of co-occurrence. These results point towards genuinely grammatical origin of sMMN response.

Syntactic MMN has been observed in non-native speakers, as well. Hanna, Shtyrov, Williams, & Pulvermüller (2016) used sMMN to assess L2 acquisition and compared native speakers' responses to those elicited from two groups of non-native speakers. More proficient non-native speakers showed native-like reduction in MMN response to syntactically valid word sequences compared to syntactically anomalous sequences, while less proficient non-native speakers' MMN responses to both conditions did not differ significantly. Consequently, the authors reason that similar brain mechanisms are used at least for some aspects of L1 and L2 grammars.

To summarise, linguistic MMN appears to be a robust and versatile indicator of early semi-automatic[7] ignition of various memory traces. Distinct lMMN and sMMN patterns have been discovered and corroborated by a number of studies, providing evidence from different languages that different types of long-term memory traces exist and interact in ways that are, at least to some extent, penetrable by the passive oddball paradigm. The most common view is that, a) strongly interconnected neural circuits, representing whole-word storage of lexical units, ignite when triggered by auditory input that matches the form of a known word, thus increasing the negativity of lMMN, and b), that mental representations of grammatically relevant units, combined by following a syntactic rule, prime each other and reduce the negativity of sMMN. Early processing of linguistic stimuli that trigger some of those memory traces can be further probed with the help of precise experimental designs and well-balanced stimuli. One of the objectives of the current thesis is exactly that: to apply the passive oddball paradigm to investigate a combination of interacting factors, namely, lexical, grammatical, and phonological.

---

[7] Even though a lot of MMN literature calls MMN an index of automatic processes, 'semi-automatic' is a more appropriate term, according to Pulvermüller & Shtyrov (2006) because the amplitude of MMN responses increases with active attention to stimuli.

3. Method

This chapter gives an overview of the method and motivates the methodological choices. Section 3.1. describes how the experiment was created, how its design is connected to the hypothesis, and what predictions it is designed to test. Section 3.2., describes how the stimuli were created, presents the data on the study subjects and lists the details about the experiment procedure. Section 3.3. gives an overview of the steps that were taken during the data processing and analysis.

3.1. Experimental design

The current study focuses on nouns with monomorphemic roots, more specifically, the noun root *krok* 'hook'. Two inflected forms of this noun are compared in the experiment: *krok-en* ($_{ACC1}$hook-SG.DEF 'the hook') and *krok-ar* ($_{ACC2}$hook-PL.INDEF 'hooks'). This specific noun root was chosen because the two inflected forms of *krok* have very similar surface frequency values[8], thus mitigating the unwanted MMN amplitude modulation of this kind.

The experiment was designed in four blocks, each with different pairs of standard and deviant stimuli (see Table 1). Each block contained 850 standard stimuli and 150 deviant stimuli. In addition, at the beginning of each block 20 "warm-up" standard stimuli were included. EEG responses to "warm-up" stimuli were not used in the data analysis.

Comparison	Block	Stimuli	Distribution
Accent 1 Enhanced lexical MMN expected	Block I	$_{acc1}$***krok-en***	15% (deviant)
		*$_{acc1}$*krok-ar*	85% (standard)
	Block II	*$_{acc1}$***krok-ar***	15% (deviant)
		$_{acc1}$*krok-en*	85% (standard)
Accent 2 Reduced syntactic MMN expected	Block III	$_{acc2}$***krok-ar***	15% (deviant)
		*$_{acc2}$*krok-en*	85% (standard)
	Block IV	*$_{acc2}$***krok-en***	15% (deviant)
		$_{acc2}$*krok-ar*	85% (standard)

Table 1: Experiment setup. Stimuli with incongruent combinations of tone accent and suffix are marked with asterisk. Deviant stimuli are in bold script; they are the minuends of an MMN curve. Standard stimuli are the subtrahends.

---

[8] I.e. 40 instances of *kroken* vs 42 instances of *krokar* in PAROLE (1997), a 24-million word corpus of Swedish language, available at https://spraakbanken.gu.se/eng/resource/parole

Each block corresponds to a separate experimental setting. Two of them target accent 1, the other two – accent 2. There are four classes of binary experimental variables that are controlled for in this setup: tone accent (accent 1, accent 2), suffix type (singular definite *-en*, plural indefinite *-ar*), stimulus rarity (standard, deviant), and stimulus validity (congruent, *incongruent).

Block I is designed to elicit MMN responses to the congruent combination of accent 1 and the singular definite suffix, i.e. a real word. Block II elicits MMN responses to the incongruent combination of accent 1 and the plural indefinite suffix, i.e. an invalid word. The goal is to measure the difference between these two MMN responses and see whether this difference is characteristic of lexically enhanced MMN. This difference is arrived at by subtracting the block II MMN curve from the block I MMN curve:

(1)   (deviant[$_{acc1}$*krok-en*] – standard[*$_{acc1}$*krok-ar*]) – (deviant[*$_{acc1}$*krok-ar*] – standard[$_{acc1}$*krok-en*])

For brevity and simplicity, equation (1) will be referred to as **lexical comparison**. Lexical comparison will test a prediction derived from the first part of the current hypothesis, namely, that deviant presentations of $_{acc1}$*krok-en* will elicit greater MMN amplitude relative to MMN amplitudes elicited by deviant presentations of *$_{acc1}$*krok-ar*.

Following the same line of reasoning, the remaining two blocks focus on combinations of accent 2 and the same suffixes. Block III elicits MMN responses to the congruent combination of accent 2 and the plural indefinite suffix, i.e. a real word. Block IV elicits MMN responses to the incongruent combination of accent 2 and the singular definite suffix, i.e. an invalid word. The difference between MMN responses to blocks III and IV is expected to be characteristic of reduced syntactic MMN. However, the order of subtraction is opposite to the one involving accent 1, so block III MMN curve is subtracted from block IV MMN curve, but not vice versa[9]:

(2)   (deviant[*$_{acc2}$*krok-en*] – standard[$_{acc2}$*krok-ar*]) – (deviant[$_{acc2}$*krok-ar*] – standard[*$_{acc2}$*krok-en*])

Equation (2) will be referred to as **syntactic comparison**. Syntactic comparison will test a prediction derived from the second part of the current hypothesis, namely, that deviant

---

[9] It is important to do the subtraction in the correct order to prevent that, where mismatch negativity manifests itself, a false positivity is found instead. Deviants that are real words with accent 1 are hypothesised to be stored in whole-word form and therefore elicit greater MMN amplitudes than invalid deviants with accent 1. Hence, for the accent 1 comparison, it is MMN responses to *invalid* deviants that ought to be subtracted from MMN responses to *valid* deviants. Meanwhile, for accent 2, the subtraction order is opposite, because deviants that are real words with accent 2 are hypothesised to be processed in a combinatorial manner involving priming effects which reduce the MMN amplitude.

presentations of $_{acc2}$*krok-ar* will elicit lesser MMN amplitude relative to MMN amplitudes elicited by deviant presentations of *$_{acc2}$*krok-en.*

To distinguish from the first-order MMN curves obtained within each block, the curves computed in lexical and syntactic comparisons will also be referred to as higher-order MMN curves.

3.2. Stimuli, subjects, and procedure

Spoken material for the stimuli was recorded in an anechoic chamber. A male Central Swedish speaker read the words without focal prominence, embedded in a carrier sentence. Since MMN response relies on any discriminable auditory changes (Näätänen, 2001), it was important to ensure that the acoustic input is the same until the point of divergence of the stimuli, i.e. suffix onset. To achieve that, the recorded stems and suffixes were measured for length, phoneme quality, and fundamental frequency in order to select the most similar ones (see Fig. 1). Selected morphemes were cross-spliced using Praat software (Boersma & Weenink, 2015) to create the congruent and incongruent combinations of stem tone and suffix. In total, four acoustically unique stimuli were created, two of them – actual words, occurring in Central Swedish, while the other two – invalid in the sense that the tone accent on the stem was incongruent with the suffix that followed (see Fig. 2). All stems were 457 ms long, all suffixes were 224 ms long. Total duration of each stimulus was 681 ms, with discrimination point (suffix onset time) at 457 ms; voicing began at 71 ms (tone onset time).
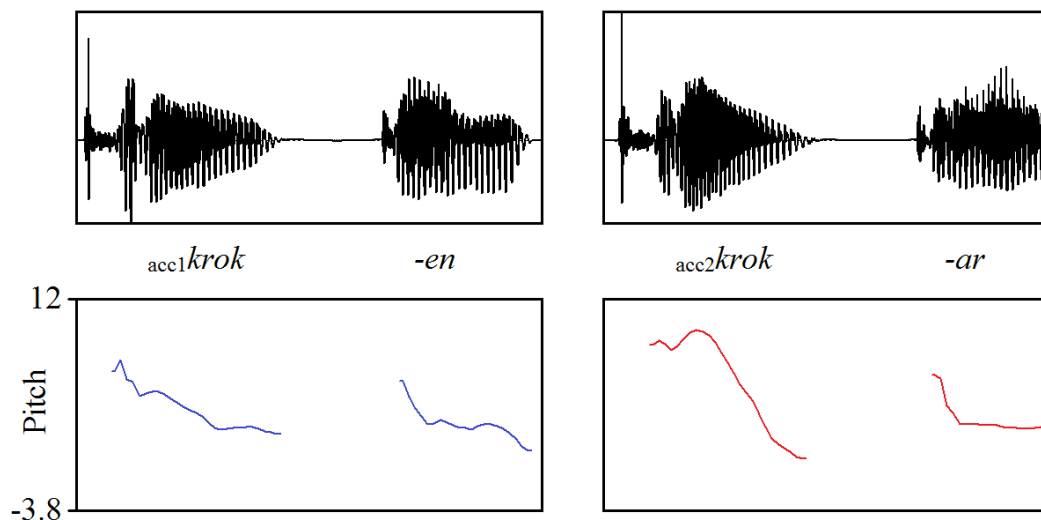

Figure 1. Acoustic waveforms and pitch curves of stimuli components. $_{acc1}$*krok-en* and $_{acc2}$*krok-ar* are two valid combinations of tone accent and suffix that represent all four building blocks used in cross-splicing procedure. Pitch is mapped on a semitone scale with reference at 100 Hz.
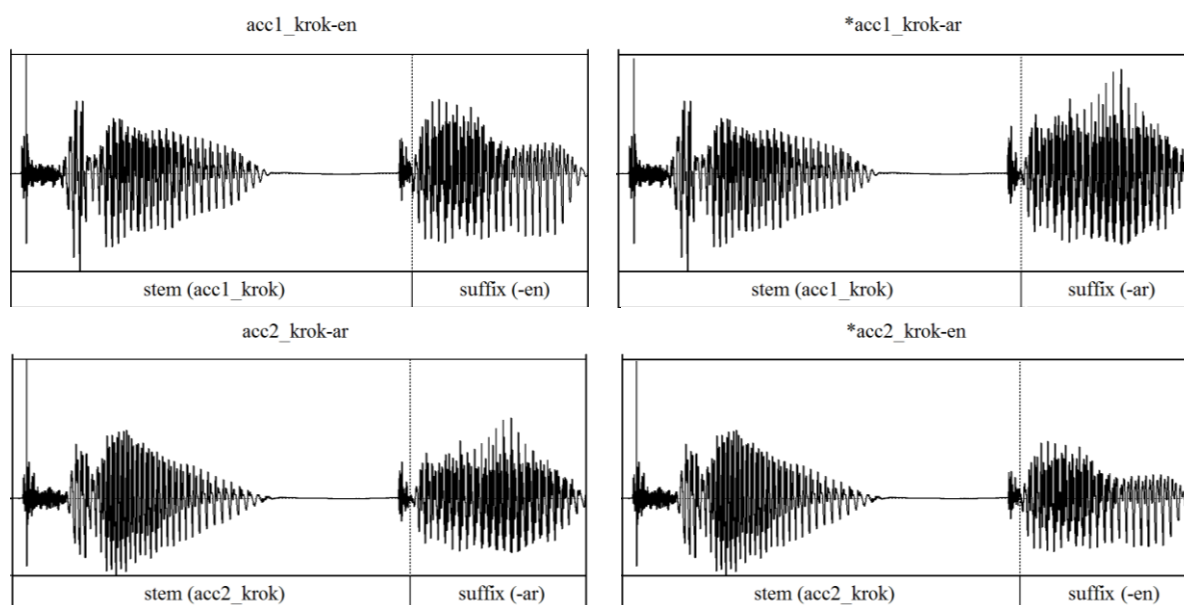
Figure 2. Acoustic and morpophonological composition of the final stimuli. Top: stimuli for lexical comparison (blocks I and II); bottom: stimuli for syntactic comparison (blocks III and IV). Point of divergence is marked by a vertical line.

Nineteen healthy, right-handed (mean handedness index 92.35R, range 65R-100R according to Edinburgh handedness questionnaire (Oldfield, 1971)) native speakers of Central Swedish (age 18-37, median age 20, 6 males) were recruited to participate in the study. Central Swedish is spoken primarily in Stockholm, Uppsala, Södermanland, and Örebro counties, so having one of these areas as their origin was one of the criteria for selection of study subjects. Before participation, the subjects gave informed consent in accordance with a protocol approved by Swedish Ethical Review Authority. Each subject participated in one experiment session, during which all 4 blocks were presented. The recording sessions lasted up to 2 hours. After the experiment, the subjects were paid for participation.

Two subjects' data were not included in the analysis. One subject revealed during the recording that they are bilingual and have acquired another language earlier than Swedish. The other subject had to terminate the experiment, so the number of trials per block was particularly unbalanced.

The subjects were seated in an electrically shielded laboratory room and watched a silent film of their own choice. The lights in the room were dimmed during the presentation of the stimuli blocks. The subjects were instructed to ignore the auditory stimuli and focus on the film. The stimuli were presented binaurally, at a comfortable volume. Stimulus onset

asynchrony was set to 1300 ms and jittered ± 50 ms in 10 ms steps[10]. PsychoPy2 application (Peirce et al., 2019) was used for presentation of stimuli. For every stimulus, three event time-stamps were made in the EEG recording, marking the onset of word, tone, and suffix. The stimuli that followed the 20 "warm-up" standards were shuffled in a semi-random way, ensuring that a deviant stimulus was always preceded by at least two standard stimuli. Each standard stimulus that followed directly after a deviant was excluded from the analysis because such standards also elicit an MMN response (Sams, Alho, & Näätänen, 1984). The order of the blocks was randomised among the study subjects to compensate for the habituation effect, i.e. a decrease in amplitude of neural response as a result of repeated presentation of the same stimuli (Sokolov, 1963). Between the blocks, short breaks were made to keep the subjects more comfortable and to minimise the habituation effect (McGee et al., 2001). During breaks, the lights in the room were brightened and the subjects could stretch their muscles, have refreshments and snacks, and switch to a different film if the first one was not engaging.

During the presentation of the auditory stimuli, electrical activity of each subject's brain was recorded with a 64-channel setup (Neuroscan system), using an EasyCap electrode cap of appropriate size, with Ag/AgCl ring electrodes mounted in an extended 10-20 system (Jasper, 1958). 2 stand-alone mastoid electrodes were used to record baseline electrical activity, for offline re-referencing. 2 bipolar electrooculogram electrodes were used to record horizontal and vertical eye movements, for enhanced compensation of ocular artefacts. The impedance at each electrode was ≤ 5 kΩ at the start of each recording session. The signal was sampled at 500 Hz and band-pass filtered online between 0.05 and 200 Hz. The CPz electrode was used as reference during the recording.

3.3. Data processing and analysis

Offline pre-processing and further processing of the recorded neurophysiological data was done in EEGLAB (Delorme & Makeig, 2004). First, a windowed sinc FIR filter was applied at transition bandwidth of 7.5 Hz and cut-off frequency of 30 Hz (-6dB half-amplitude), using Hamming window. Then, the data was re-referenced against the average of both mastoids. A visual inspection of the recorded brainwaves was carried out and sections with particularly powerful artefacts were cut out and discarded. After manual pruning, independent component

---

[10] With regular stimulation, alpha waves (brainwaves that are dominant when a subject is resting, in a calm, alert state) may synchronise with the rhythm of stimulus presentation thus introducing noise that is inseparable from data; jittering of stimulus onset times helps to minimise this effect (Luck, 2014).

analysis (ICA) was run separately on each subject's data. ADJUST plugin (Mognon, Jovicich, Bruzzone, & Buiatti, 2011) and IC label tutorial (Pion-Tonachini, 2019) were used to evaluate the results of ICA and reject the eye, muscle, and channel noise components from the data. Later on, ERPLAB plugin (Lopez-Calderon & Luck, 2014) was used to segment the continuous data into epochs and process them. 800 ms long epochs were extracted, with epoch boundaries set between -200 ms and 600 ms relative to suffix onset time. Outlier epochs whose signal amplitude exceeded ± 100 μV, were removed from averaging and further analysis. Baseline correction was done on the 200 ms time window preceding the suffix onset. Then, average event-related potentials were calculated for each subject, stimulus, and electrode. Due to individual amounts of heavy artefacts and varying numbers of epochs that needed to be rejected, the number of remaining trials per subject per stimulus was not equal. In order to compensate for this variation, a weighted averaging procedure was applied. Afterwards, the MMN curves were calculated for each block.

MMN responses have their maximal amplitude over frontal and central recording sites (Näätänen & Winkler, 1999). Therefore, visual examination of the averaged data started at FCz and Cz electrodes. The first-order MMN curves were paired by accent conditions and overlaid on top of each other, to show the difference between MMN responses at these recording sites.

First, the effects were localised temporally by identifying the negative and positive peaks and their local latencies in the MMN curves. Around these peaks, 40-ms-wide time windows were defined for statistical analysis. Spatial localisation of the effects was done by visually inspecting the topographical maps that show how the electrical activity is distributed over the scalp. Narrow regions of interest (ROI) were defined separately for each time window. Average electrical activity throughout the given time window was calculated and plotted as a topographical map of the scalp. Then, a cluster of electrodes with greatest activation amplitudes was selected for statistical analysis as the ROI for that time window. The cluster size was kept at or below 6 electrodes because small ROIs have an advantage over larger ROIs in detecting true within-region effects (Pataky, Robinson, & Vanrenterghem, 2016: 7). Topographical maps were also used as a guide when identifying MMN-like activity − widespread negative deflection in frontocentral sites − and differentiating it from other kinds of event-related responses.

Paired-samples t-tests were run for lexical and syntactic comparisons, to check if there were any statistically significant differences within the designated time windows and ROIs.

1-tailed tests were conducted because the hypothesis predicts negative deflections, effectively putting the positive tail of the distribution outside the significance range.

An MMN curve is a result of a subtraction and, as such, it does not inform about exact contributions by its minuend relative to its subtrahend. These contributions are of great interest though, especially because the empirical object of the current study involves an added dimension – tone accent. While a lot of similar MMN studies have been focusing on non-tonal languages, such as English (Pulvermüller & Shtyrov, 2003), Finnish (Pulvermüller, Shtyrov, Ilmoniemi, & Marslen-Wilson, 2006), French (Hanna et al., 2014), or German (Pulvermüller & Assadollahi, 2007), and dealing with interaction of lexical, morphological, and syntactic processing, the current study aims, in addition, to deal with the phonological aspect of the morphemes or lexemes triggering those processes. Therefore, to examine which (if any) contributions of the different types of stimuli were significant, complementary 2-tailed one-sample t-tests were run for each block.

Lastly, visual inspection of the ERP waveforms elicited in each stimuli condition was carried out in order to assess the neurophysiological response differences across the experimental variables tone accent, suffix type, and stimulus rarity. However, no statistical analysis was done to investigate these effects because the main focus of this thesis is on the MMN effects.

4. Results

This chapter presents a summary of the results and how they were obtained. The analysis consists of three parts, which is reflected in the structure of this chapter. In 4.1., results of the main statistical analysis of the lexical and syntactic MMN comparisons are described. In 4.2., results of the complementary statistical analysis of blockwise contributions to the MMN curves are described. In 4.3., results of the visual analysis of the ERP waves are presented in light of effects attributable to the different experimental variables.

4.1. MMN comparisons

MMN curves from all four blocks were extracted at FCz electrode and aligned into lexical and syntactic comparisons (Fig. 3). Higher-order MMN curves were computed (Fig. 4).
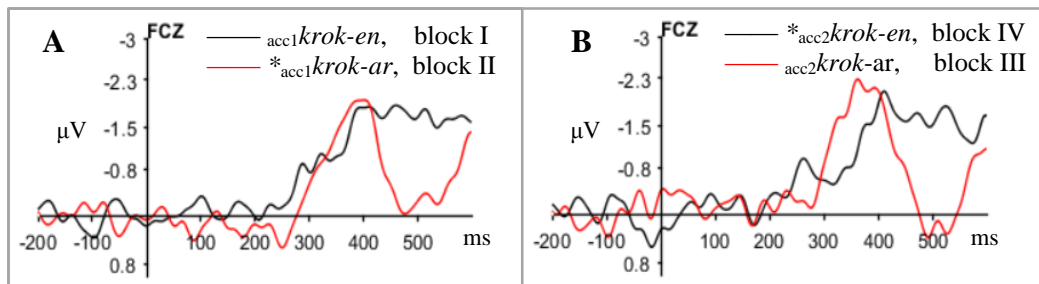

Figure 3. First-order MMN curves showing the MMN responses to deviants in separate blocks. A: Blocks I and II, used in the lexical comparison. B: Blocks III and IV, used in the syntactic comparison.
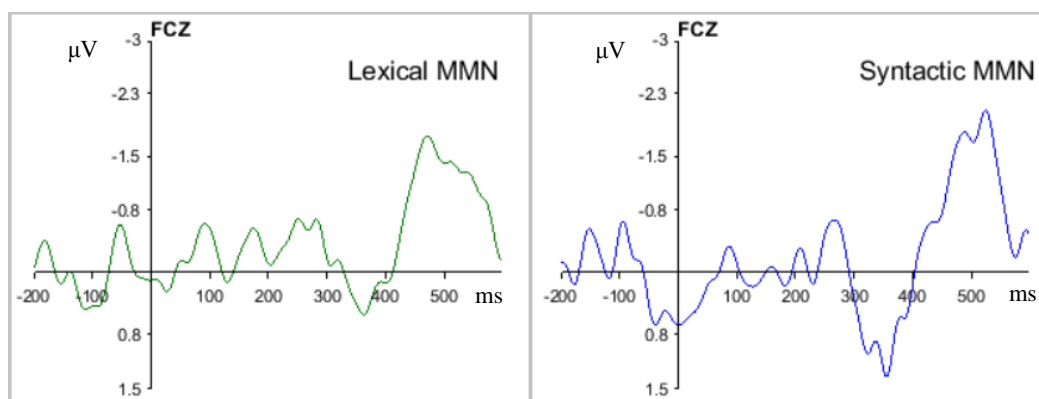

Figure 4. Higher-order MMN curves. Lexical MMN curve represents block II MMN subtracted from block I MMN. Syntactic MMN curve represents block III MMN subtracted from block IV MMN.

Three negative peaks within MMN latency range were elicited in the lexical comparison. In the syntactic comparison, three negative peaks were also elicited at roughly the same latencies, but only the third peak was of comparable magnitude relative to lexical comparison. The early two negative peaks were smaller and lasted for shorter time. The middle negative peak in the syntactic comparison was also later (~200 ms) than the middle negative peak in the lexical comparison (~175 ms). In both conditions, the MMN-like ERP effects were followed by a positive deflection peaking at ~350 ms, this positivity was much more pronounced in the syntactic comparison. Afterwards, a great negative-going wave was observed in both conditions. It was wider in the lexical comparison, peaking at ~475 ms. In the syntactic comparison, it peaked later, at ~515 ms but had a slightly narrower and steeper waveform. The waveforms of the negative peaks in the MMN latency range were also narrower and steeper in the syntactic comparison.

Based on exact measurements (Table 2) of the negative deflections observed in both comparisons, common time windows were defined for statistical analysis of ERP effects within MMN latency range: 70-110 ms (very early MMN), 150-190 ms (early MMN), 255-295 ms (late MMN)[11].

	Peak latency and amplitude at FCz		Common time window [mean peak latency]
	Lexical comparison	Syntactic comparison	
1st peak	92 ms, –0.632 μV	88 ms, –0.334 μV	70-110 ms [90 ms]
2nd peak	175 ms, –0.568 μV	160 ms, –0.074 μV	150-190 ms [167.5 ms]
3rd peak	282 ms, –0.69 μV	267 ms, –0.67 μV	255-295 ms [274.5 ms]

Table 2. Common MMN time windows, based on individual and mean values for peak amplitude and latency of the negative deflections in lexical and syntactic comparisons within MMN time range.

Scalp distributions for the negative deflections in each time window were inspected (Table 3 on next page) and used to define the regions of interest (ROI) for statistical analysis.

---

[11] The labels very early, early and late MMN were adopted from Pulvermüller & Shtyrov (2003) because they found negative MMN deflections spread in similar latencies.

ERP / time window	Lexical comparison	Syntactic comparison
Very early MMN / 70-110 ms	Bin15,70-110 ms	Bin14,70-110 ms
Early MMN / 150-190 ms	Bin16,150-190 ms	Bin14,150-190 ms
Late MMN / 255-295 ms	Bin16,255-295 ms	Bin14,255-295 ms
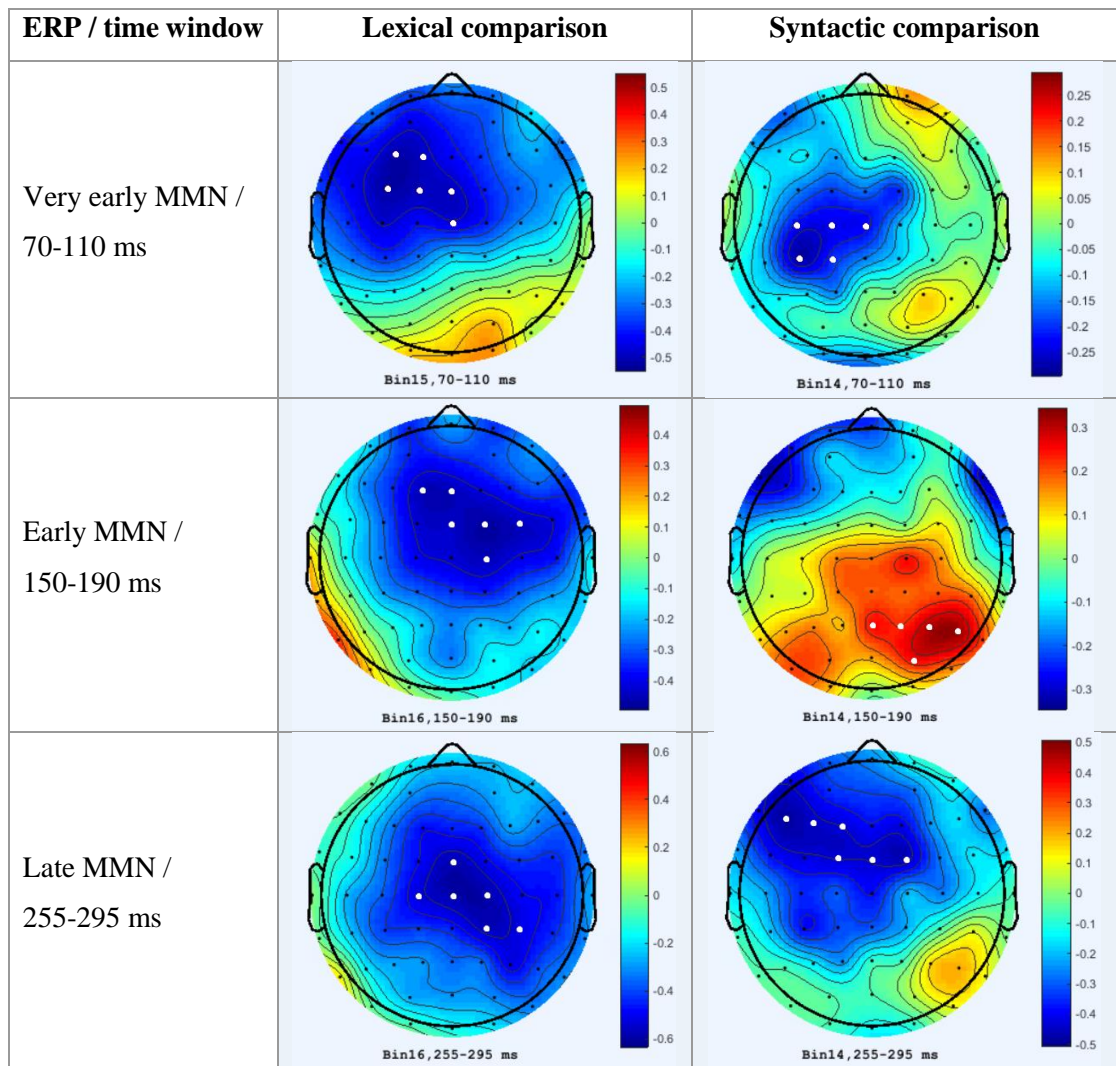
Table 3. Mean scalp topographies for lexical and syntactic comparisons over selected time windows. White dots represent electrodes that constitute the ROI for statistical analysis.

1-tailed paired-samples t-tests were conducted to compare the MMN amplitudes between blocks I ($_{acc1}$*krok-en*) and II (*$_{acc1}$*krok-ar*) in the lexical comparison, and between blocks IV (*$_{acc2}$*krok-en*) and III ($_{acc2}$*krok-ar*) in the syntactic comparison. The results are summarised in Table 4.

ERP / Time window	Lexical comparison		Syntactic comparison	
	$_{acc1}$*krok-en*	$*_{acc1}$*krok-ar*	$_{acc2}$*krok-ar*	$*_{acc2}$*krok-en*
Very early MMN / 70-110 ms	M = −0.22 SD = 0.93	M = 0.27 SD = 0.68	M = −0.18 SD = 0.83	M = −0.37 SD = 0.73
	p = **0.036** t = −1.926		p = 0.284 t = 0.584	
Early MMN / 150-190 ms	M = −0.11 SD = 0.81	M = 0.34 SD = 0.6	M = −0.03 SD = 0.93	M = −0.22 SD = 0.72
	p = **0.045** t = −1.803		p = 0.214 t = 0.815	
Late MMN / 255-295 ms	M = −0.41 SD = 0.85	M = 0.13 SD = 0.95	M = −0.31 SD = 0.98	M = −0.74 SD = 1.03
	p = **0.027** t = −2.072		p = 0.124 t = 1.198	

Table 4. Main MMN findings. Results of 1-tailed paired-samples t-tests comparing the MMN curves in lexical and syntactic conditions. P-values over 95% confidence interval are bolded.

In the lexical comparison, $_{acc1}$*krok-en* elicited a **very early MMN** peak at 101 ms; the effect was significant in the left anterior recording sites compared to $*_{acc1}$*krok-ar* in the time window of 70-110 ms, t(16) = −1.926, p = 0.036. Later, $_{acc1}$*krok-en* elicited an **early MMN** peak at 169 ms; the effect was significant in the frontocentral and centre-right recording sites compared to $*_{acc1}$*krok-ar* in the time window of 150-190 ms, t(16) = −1.803, p = 0.045. After this, $_{acc1}$*krok-en* elicited a **late MMN** peak at 287 ms; the effect was significant in the central recording sites in the time window of 255-295 ms, t(16) = −2.072, p = 0.027. None of the negative peaks in syntactic comparison showed statistically significant results.

4.2.Analysis of ERP waves within blocks

The above tests were conducted for the higher-order MMN curves laid out as lexical and syntactic comparisons. However, as mentioned in the Method section, an MMN curve is a result of a subtraction, and is not transparent when it comes to individual contributions of its minuend and subtrahend. These contributions, however, are of direct interest in order to better understand what processes were potentially taking place during the recordings. To gain some insight into the complexity of the effects that were discovered on the level of lexical and syntactic comparisons, a visual analysis was carried out on the 'building blocks' of the first-order MMN curves, i.e. the ERP waves elicited by deviant and standard stimuli in each block. These responses were averaged across the electrodes that made up the ROIs defined previously. Within-block comparisons at appropriate ROIs were plotted (see Fig. 5 and 7) and

relevant time windows were highlighted. Then, 2-tailed one-sample t-tests were conducted, comparing the differences within each block to a null hypothesis (which stated that, within a given block, there was no difference between responses to deviants and to standards). Significant p values that were found during these tests are included in Figures 5 and 7; for complete results of these tests, see the Appendix.
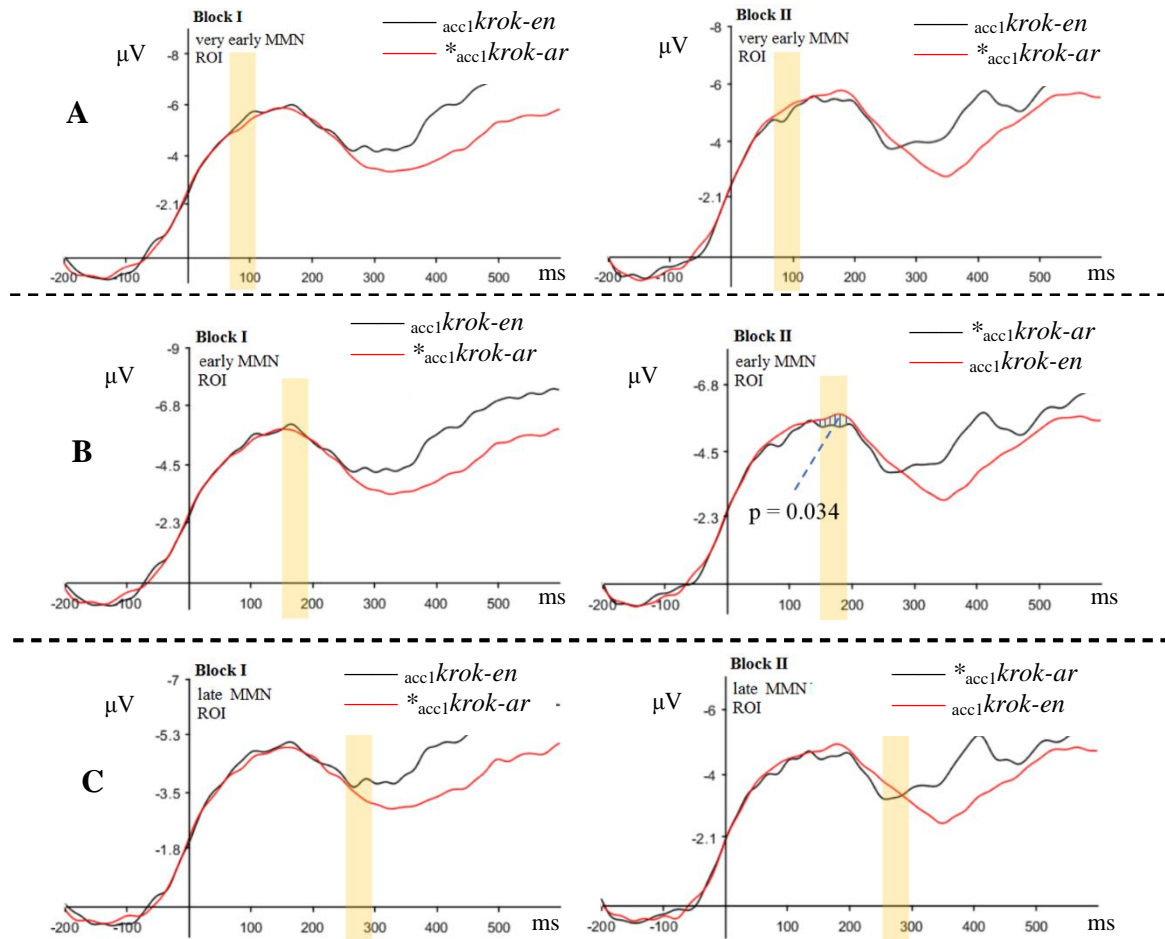

Figure 5. ERP comparisons within blocks I and II. Responses to deviant stimuli are traced in black, those to standard stimuli – in red. Response waves are averaged across the ROI of the appropriate time window. Time windows of the relevant ERPs are marked with yellow bars. **A**: Very early MMN across electrodes Cz, FCz, FC1, F1, FC3 and F3. **B**: Early MMN across electrodes F1, Fz, FCz, FC2, C2 and FC4. **C**: Late MMN across electrodes Cz, FCz, C2, CP2, C1 and CP4.

Contributions to lexical comparison. Visual inspection revealed that, during the very early MMN and early MMN time windows, responses to deviant $_{acc1}$*krok-en* in block I contributed to the MMN curve with increased negativity, whereas responses to deviant *$_{acc1}$*krok-ar* in block II contributed with decreased negativity. Subtracting the latter from the former, both contributions are in effect adding up to an enhanced MMN. Compared to null hypothesis, the peak observed during the early MMN time window in the MMN curve of

block II was statistically significant; t = 2.315, p = 0.034. To summarise, the ERP waves behaved in line with the hypothesis – a real word enhanced the negativity, while the incongruent word reduced it. Thus, the very early MMN and early MMN effects appear to be true positive findings.

Conversely, the peak in the lexical comparison that was identified as late MMN appears to be a coincidence and not an actual effect. None of the ERP waves in both blocks showed any peaks confined to that time window. Instead, the peak in the higher-order MMN curve seems to be a combination of the onset of a larger and prolonged negativity elicited by deviant $_{acc1}krok\text{-}en$ in block I and the offset of the positive deflection elicited by deviant $*_{acc1}krok\text{-}ar$ in block II. A large, prolonged negativity has in fact been elicited by all deviants from around 300 ms onwards (see subsection 4.3. about this and other ERP effects) and is so great that its onset alone may be sufficient to result in p<0.05 that was detected by the initial statistical test in the late MMN window. It appeared that an actual late MMN effect may be there in block I (see Fig. 6C, a negative deflection in block I at around 210-250 ms), only too early to be captured by the common time window which was set to 255-295 ms. However, the mean topographical map of block I MMN curve at 210-250 ms (Fig. 6) revealed that the negativity has a centre-right distribution that is not typical of MMN. Hence, the statistically significant results for late MMN in lexical comparison can be deemed a false positive.
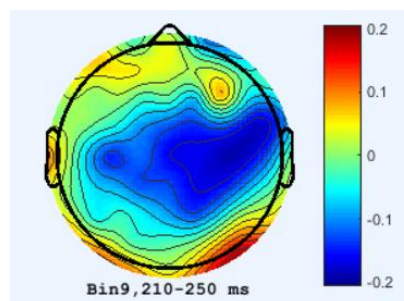
Figure 6. Mean scalp topography of block I MMN curve during late MMN time window.

Contributions to syntactic comparison. The overall visual impression is that responses constituting the syntactic comparison within MMN latency range (around 70-300 ms) clearly differed from those in the lexical comparison. Responses to deviants and standards in blocks I and II are closely aligned and overlapped most of the time, whereas blocks III and IV display a rather steady difference in the amplitudes of responses to deviants relative to standards. Also, the response dynamics differs with respect to polarity. In syntactic comparison deviants of blocks III and IV both elicited more negative responses than their respective standards. In lexical comparison, on the other hand, deviants of block I elicited more negative responses

relative to standards, while deviants of block II elicited less negative responses relative to standards.

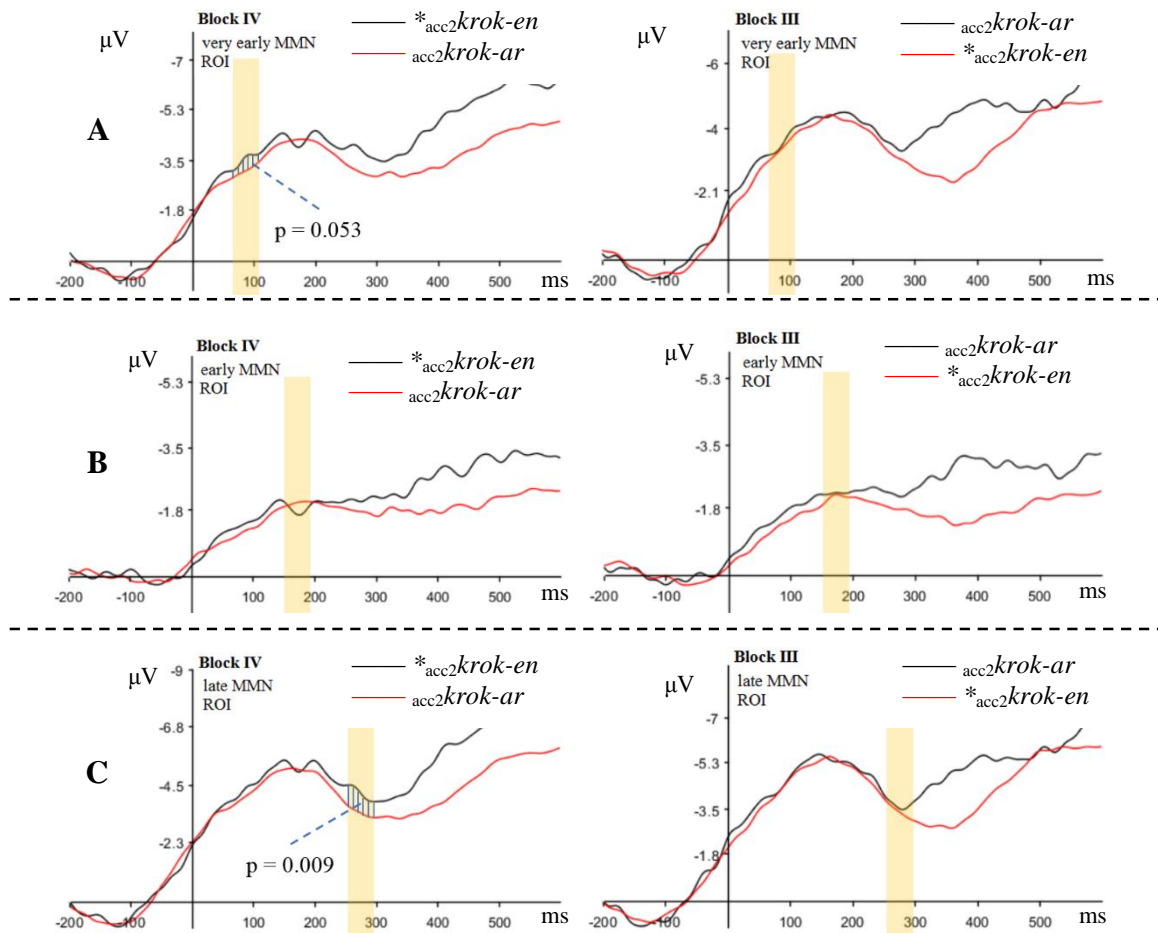**Figure 7.** ERP comparisons within blocks IV and III. Responses to deviant stimuli are traced in black, those to standard stimuli – in red. Response waves are averaged across the ROI of the appropriate time window. Time windows of the relevant ERPs are marked with yellow bars. **A**: Very early MMN across electrodes CP3, C3, C1, Cz and CP1. **B**: Early MMN across electrodes P4, P6, Pz, P2 and PO4. **C**: Late MMN across electrodes F5, F3, F1, FC1, FCz and FC2.

Focusing on the syntactic comparison, a tendency is seen during all three MMN time windows where, in block III, the deviant ERP wave approached the standard ERP wave, while in block IV the deviant wave moved slightly further apart from the standard wave. Responses to deviant $_{acc2}$*krok-ar* in block III contributed with a slight decrease in negativity during all three time windows, which is in line with the prediction of reduced MMN for congruent combinatorial sequences. Statistically, though, these deflections were not significant. Only contributions from block IV were significant compared to null hypothesis. During the very early MMN window, deviant *$_{acc2}$*krok-en* relative to standard in block IV elicited a negativity increase on the verge of statistical significance; $t = -2.087$, $p = 0.053$. During the late MMN window, deviant *$_{acc2}$*krok-en* relative to standard in block IV elicited a significant negativity

increase; t = –2.96, p = 0.009. However, the latter negativity increase is due to onset of a larger and more prolonged negative deflection seen in all blocks during the second half of the EEG epoch, and was therefore not interpreted as an MMN effect.

4.3.ERP waves and effects of experimental variables

Past around 300 ms after stimulus onset, clear differences between neural responses to different stimuli emerge. Primary ERP waves have been plotted (Fig. 8) and inspected in order to interpret these effects. Effects of tone accent, suffix type, and stimulus rarity are reflected by visible differences in the ERP waveforms, as described below.
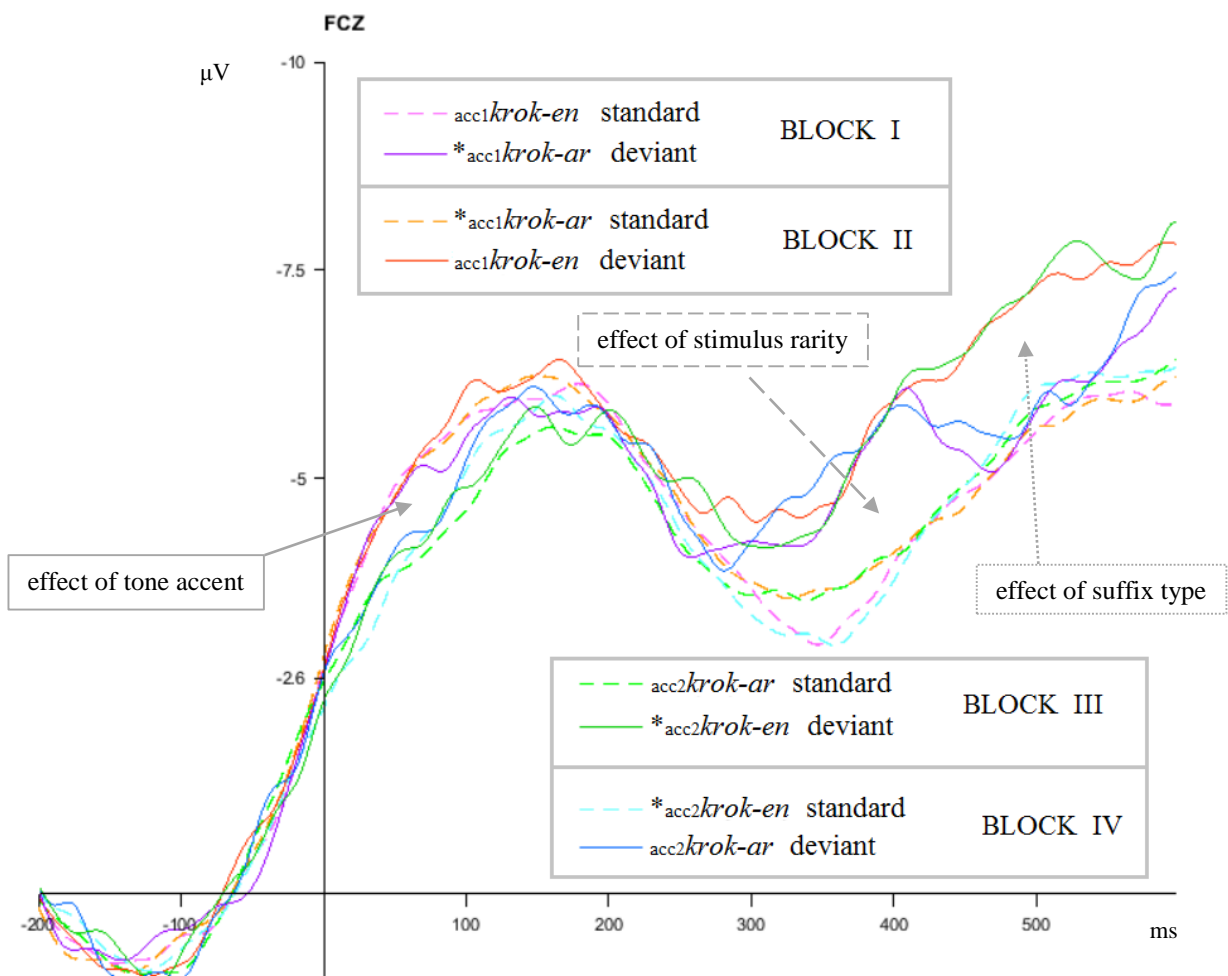
Figure 8. Average ERP waves of all stimulus conditions at FCz. Deviant presentations are traced in continuous lines, standard presentations, in dashed lines. Responses to accent 1 have warm colours (block I – pink/purple, block II – red/orange), responses to accent 2, cold colours (block III – green, block IV – blue). Arrows point at gap areas between different ERP waves and text in the boxes marks which experimental variables differentiate the stimulus conditions separated by those gaps.

Figure 8 shows that all eight stimulus conditions have elicited rather similar waveforms consisting of three main movements: an early negative peak, followed by a positive-going wave which then turned into a prominent negative deflection. The ERP waves for deviant stimuli featuring accent 2 also show a second positive movement, although less prominent than the first one.

The first negative deflection peaks at around 150 ms for accent 1 conditions, slightly later for accent 2 conditions. The ERP waves group by tone accent in the very beginning: stimuli featuring accent 1 (warm colours) elicited an earlier and somewhat larger negative deflection than stimuli featuring accent 2 (cold colours). In conjunction with frontocentral distribution consistent through all eight stimulus conditions, this effect looks like pre-activation negativity for accent 1 words.

After the first negative peak, a positive-going movement takes place in all conditions. Notably, the latencies when responses from different stimulus conditions reach their turning point back to negative-going movement are more spread than the peak latencies of the initial negative wave. Also, tone accent, suffix type, and stimulus rarity all seem to modulate the positive movement's amplitude and latency, but to different extent. The stimuli presented as standards elicited a greater and longer decrease in negativity with a turning point at around 350 ms, compared to stimuli presented as deviants where the negativity starts increasing again already at around 300 ms. The positive movement elicited by standards with singular definite suffix (dashed pink and dashed light blue lines) goes further down than the one elicited by standards with plural indefinite suffix (dashed green and dashed orange lines). Among deviants, the suffix type also seems to have played a role: deviants with the singular definite suffix showed a faster and steeper positive movement than those with plural indefinite suffix.

Lastly, a great negative-going wave is seen in all stimulus conditions past 350 ms after stimulus onset; apparently, it peaked either around or past the epoch cut-off time. For deviants, compared to standards, this negativity shows an earlier onset, greater amplitude, and, interestingly, an effect of the suffix type. The singular definite suffix elicited a great negativity that continued to increase all the way until the end of the epoch window, whereas the ERP waves of stimuli with the plural indefinite suffix suddenly changed course and displayed a positive-going wave just after 400 ms after stimulus onset, peaking at around 480-500 ms. For standards, in the meantime, this negative wave is rather uniform, regardless of other experimental variables.

5. Discussion

This chapter discusses the results presented in the previous chapter. Section 5.1. interprets the MMN results by integrating the findings from main and complementary statistical analysis. It also situates the findings in context of similar studies of tone accents. Section 5.2. features a discussion of the ERP results, while Section 5.3. interprets the important ERP findings in light of the hypothesis.

5.1.MMN comparisons

Results of the lexical comparison corroborate the first prediction of the hypothesis: the deviant $_{acc1}$*krok-en* elicited statistically significant negativity increase in all three MMN time windows compared to negativity levels elicited by the deviant *$_{acc1}$*krok-ar*. This supports the claim that valid accent 1 words are retrieved from lexical memory in whole-word forms. Results of the syntactic comparison are less straightforward; the differences between deviant *$_{acc2}$*krok-en* and deviant $_{acc2}$*krok-ar* did not turn out to be statistically significant in any of the defined MMN time windows. However, the t-values found for syntactic comparison are all positive, suggesting consistently reduced MMN responses (cf. negative t-values found for the enhanced MMN in lexical comparison). Hence, the second prediction of the hypothesis is not falsified by these results.

Results of the within-block analysis corroborated the very early MMN and early MMN effects found in the lexical comparison but not the late MMN effect. Regarding the syntactic comparison, within-block analysis supported the tendency of reduced MMN response for $_{acc2}$*krok-ar* which is in line with the second prediction of the hypothesis. To summarise, the lMMN was clearer than sMMN, which suggests that accent 2 words are not necessarily processed strictly in accordance with morphological rules. To interpret these results, it is useful to consider what is known about the sources of processing of tone accents. Schremm et al. (2018) found a correlation between cortical thickness in left planum temporale and processing speed for real words, regardless of tone accent, which they interpreted as an indication of whole-word retrieval. This interpretation supports the idea that accent 2 words may also be stored in whole-word representations in the brain. Since the incongruent stem+suffix combinations used in the current study contained an actual root morpheme that exists in Swedish lexicon, less clear sMMN for these combinations might indicate that such memory traces were indeed activated, albeit to a lesser degree because accent 2 words still

elicited some tendencies supportive of morphological decomposition processes. One explanation of such an activation pattern for accent 2 conditions could be that the neural circuits serving lexical, whole-word retrieval and those involved in morphological, rule-based decomposition are both valid processing methods for accent 2 words and can be initiated in parallel. Then, the lesser but not non-existent sMMN might signal a real-time competition between the two processing routes. This explanation is compatible with the findings by Söderström, Horne, Mannfolk, et al. (2017). They found that the brain region associated with suppression of lexical candidates (pars opercularis of the inferior frontal gyrus) is involved during processing of accent 2 words. Considering this, it could be possible that when an accent 2 word is heard, lexical pre-activation takes place first and, when the processing mechanisms detect that the amount of pre-activated continuations is inefficiently big, the lexical continuation candidates are suppressed and instead a morphological rule is invoked to aid the recognition and integration process.

Results of the current investigation also resemble those presented by Zora et al. (2019). They found that the lMMN elicited by lexically stressed stems was larger than the sMMN elicited by phonologically stressed stems. In light of these findings, it seems reasonable to assume a possibility of a functional connection between the processes behind post-lexical assignment of tone accent and those underlying phonological (i.e. post-lexical) assignment of stress. Then, one may speculate that lexical and combinatorial processing routes are initiated in parallel not only by accent 2 words but by other post-lexically defined word forms, too. However, the possibility of tone accent and stress categories being related on the neural level should be investigated further, before any claims can be made.

5.2.ERP waves and effects of experimental variables

The section above summarised and discussed the early effects of stimulus validity (because that is what the MMN response reflects). At later latencies, looking back at the first-order MMN curves (Fig. 3), great differences in amplitude are seen past 250 ms after stimulus onset. However, this should not be interpreted as effects of stimulus validity. Looking at the ERP waves (Fig. 8) it becomes apparent that these late effects are mostly due to suffix type and stimulus rarity. Therefore, the true effects of stimulus validity at those latencies, if any, would be covered up and, unfortunately, it is difficult to disentangle these variables in an MMN study design. On the other hand, it is plausible that there are no important effects related to stimulus validity at the later latencies at all. The type of rule that the invalid stimuli violated is

phonological and since phonological processing takes place early, around 100 ms after stimulus onset (Friederici, 2017), the MMN analysis can be taken as a sufficient analysis of validity effects.

Effect of tone accent manifested as the gap between responses to accent 1 and accent 2 stimuli before 100 ms after stimulus onset (Fig. 8, continuous arrow) and is interpreted as pre-activation negativity for accent 1 words. The scalp distribution (frontocentral negative maximum) and functional relation to tone accents meets the defining criteria of PrAN (Söderström et al., 2016). Admittedly, PrAN overlaps with the MMN time scope here but this should not affect the results of MMN comparisons because the stimuli within each MMN comparison share the same accent while PrAN is observed between accents.

A clear effect of stimulus rarity is seen during the negative-going deflection past 350 ms after stimulus onset (Fig. 8, dashed arrow). In general, this wave is reminiscent of N400 component. Since N400 is sensitive to semantic content of the stimuli, real words should elicit smaller N400 amplitude than pseudowords (Swaab et al., 2012) which is not the case here. Instead, it is the standards that have elicited smaller negative amplitude than deviants. One possibility is that this decrement in response is related to habituation effects caused by repeated exposure to standard stimuli (Marta Kutas & Federmeier, 2000). Another way of looking at this is to consider the element of surprise which is inherent to deviant presentations by design. Deviants stick out among standard stimuli and so, if the deviants caused sufficient surprise to attract the subjects' attention, a semi-conscious semantic processing of the word might have been initiated, resulting in N400. Returning to the fact that incongruent combinations of tone accent and suffix did not elicit a distinct N400, a quite likely explanation could be that the accent 1 word *kroken* does not constitute a minimal tonal pair (i.e. a combination of this word form with accent 2 is meaningless and does not form a valid lexeme in Swedish), so that occurrences of wrong tone accent in *kroken* would not introduce semantic ambiguities.

5.3. Type of suffix revisited

The effect of suffix type manifests as a gap between ERP waves starting from 400 ms after stimulus onset (Fig. 8, dotted arrow); here the deviants with the plural indefinite suffix *-ar* showed a distinct positive movement at 400 ms after stimulus onset, while responses to deviants with singular definite suffix *-en* continued to increase in negativity. Considering that scalp topography showed widespread frontocentral distribution, it is likely that this positive shift indicates the P600 component elicited by deviants ending with *-ar*. This component is

associated with re-evaluation of syntactic information. The interpretation of the effect of suffix type as P600 implies that the singular definite suffix is integrated into the lexical representation of the word forms, while the plural indefinite suffix is not. At this moment it is worth to remember that these two suffixes are morphologically not equivalent. The singular definite suffix can be more precisely referred to as the postposed definite article clitic (see p. 2, 2nd footnote). As pointed out by, e.g. Riad (1998), the clitic -*en* does not assign any tone accent – contrary to suffixes proper, of which the majority assign accent 2, including the plural indefinite suffix -*ar*. The fact that the morphological status of -*en* and -*ar* is not equal from the tonal perspective, may be important in light of the current findings. If the effect of suffix type has been correctly interpreted as P600, it points at systematic differences in neural processing routes triggered by accent-inducing suffixes compared to accent-neutral clitics. In other words, this effect seems to reflect a neural correlate of the tone-assigning status that differentiates the two morpheme types. A theory suggests itself that plural indefinite word forms ending with -*ar* and, perhaps, other accent 2-inducing suffixes are decomposed syntactically, whereas singular definite word forms ending with -*en* are not. A tempting explanation as to why the latter forms did not elicit a response indicative of syntactic re-evaluation in the current study could be exactly what the hypothesis predicts – that they are stored in whole-word form and are not subjected to further syntactic analysis after the word has been recognised and integrated.

From the hypothesis point of view, the possibility that clitic morphemes are stored along with the root in singular definite word forms is unexpected but welcome. On one hand, no predictions were made in the current work regarding the difference between suffix and clitic status and their neural correlates (on the contrary, it was presumed that both morpheme types can be subsumed under the suffix category for the purposes of the current investigation). On the other hand, the assumption that the accent 2-inducing suffix is processed in a combinatorial fashion while the accent-neutral clitic is stored along with the accent 1 root in whole-word form does not contradict the current hypothesis. Rather, if this assumption is valid, it would clearly support the main ideas behind the hypothesis, namely, that accent 2 is processed as a feature of suffix morphemes and accent 1 is represented holistically.

5.4. Method discussion

In the current study, source reconstruction was not carried out, which would have aided interpretation of the results. However, other studies have been carried out, e.g. those by Schremm et al. (2018) and Söderström, Horne, Mannfolk, et al. (2017), where EEG and

neuroimaging techniques such as fMRI have been employed together, so the results of the current study could still be interpreted in light of the findings and suggestions made by researchers who had data on source activation patterns and how they correlate with EEG responses.

6. Concluding remarks

6.1. Summary

The statistical analysis confirmed that word forms with accent 1 elicited significant lMMN responses, whereas word forms with accent 2 did not elicit significant sMMN response. The combinatorial nature of accent 2 words was supported by another finding. Through analysis of ERP effects it was discovered that the type of suffix morpheme affects linguistic processes in a way that provides support to the hypothesis: the accent 2-inducing plural indefinite suffix elicited an ERP component that was interpreted as P600 indicating additional syntactic processing, whereas the accent-neutral plural indefinite suffix did not. The difference between these two suffix types is that one of them is an inflectional suffix whereas the latter is a postposed definite clitic. On the whole, the combined results allow to answer both research questions positively: 1) The evidence collected in the current MMN experiment is in line with previous experiments and supports the claim that accent 1 is represented holistically in whole-word forms; 2) There is a reason to believe that accent 2 is represented in the brain as a feature of suffix morphemes. In addition, the sMMN results were interpreted as an indication of possibly parallel processing of accent 2 words that would involve both lexical, whole-word representations and morphological, rule-based, combinatorial operations. The ERP results also point at a potential existence of a neural correlate of suffix versus clitic morphemes in tone accent context.

6.2. Suggestions for future studies

It would be interesting to further explore the potential neural correlates of the suffix-clitic difference and how it interacts with the tone accent system in Swedish. First of all, an attempt to falsify or replicate the results of the current study seems in order, using a dedicated experimental design that would better suit this problem. In case of replication, experiments with other types of root morphemes could be carried out to see if the suffix-clitic distinction is sensitive to, e.g. prosodic shape of the word form, word frequency, existence of minimal tonal pairs etc.

Appendix: *Within-block t-test statistics.*

Results of 2-tailed one-sample t-tests carried out to check the significance of MMN responses within each block:

Time window	Block I $_{acc1}krok\text{-}en-$ $*_{acc1}krok\text{-}ar$	Block II $*_{acc1}krok\text{-}ar-$ $_{acc1}krok\text{-}en$	Block III $_{acc2}krok\text{-}ar-$ $*_{acc2}krok\text{-}en$	Block IV $*_{acc2}krok\text{-}en-$ $_{acc2}krok\text{-}ar$
Very early MMN	M = -0.22 SD = 0.93 t = −0,984 p = 0.34	M = 0.27 SD = 0.68 t = 1,657 p = 0.117	M = -0.18 SD = 0.83 t = -0.887 p = 0.388	M = -0.37 SD = 0.73 t = -2.087 p = 0.053
Early MMN	M = -0.11 SD = 0.81 t = 0,572 p = 0.575	M = 0.34 SD = 0.6 t = 2,315 p = 0,034	M = -0.03 SD = 0.93 t = -0.129 p = 0.899	M = 0.22 SD = 0.72 t = 1.243 p = 0.232
Late MMN	M = -0.41 SD = 0.85 t = -1.961 p = 0.068	M = 0.13 SD = 0.95 t = 0.571 p = 0.576	M = -0.31 SD = 0.98 t = -1.31 p = 0.209	M = -0.74 SD = 1.03 t = -2.96 p = 0.009

References

Alexandrov, A., Boricheva, O., Pulvermüller, F., & Shtyrov, Y. (2011). Strength of word-specific neural memory traces assessed electrophysiologically. *PLoS ONE*, *6*(8), 1-8.

Ambrazaitis, G. (2009). *Nuclear Intonation in Swedish: Evidence from Experimental-Phonetic Studies and a Comparison with German*. Lund University.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, *37*, 94–117.

Bakker, I., MacGregor, L. J., Pulvermüller, F., & Shtyrov, Y. (2013). Past tense in the brain's time: Neurophysiological evidence for dual-route processing of past-tense verbs. *NeuroImage*, *71*, 187–195.

Boersma, P., & Weenink, D. (2015). Praat: doing phonetics by computer [computer program].

Bonte, M. L., Mitterer, H., Zellagui, N., Poelmans, H., & Blomert, L. (2005). Auditory cortical tuning to statistical regularities in phonology. *Clinical Neurophysiology*, *116*, 2765-2774.

Bruce, G. (1977). *Swedish word accents in sentence perspective*. Liber.

Bruce, G. (1998). *Allmän och svensk prosodi (Praktisk lingvistik 16)*. Department of Linguistics and Phonetics, Lund University.

Bruce, G., & Gårding, E. (1981). A prosodic typology for Swedish dialects. In *Working Papers* (Vol. 21, pp. 69-75). Lund University.

Cappelle, B., Shtyrov, Y., & Pulvermüller, F. (2010). Heating up or cooling up the brain? MEG evidence that phrasal verbs are lexical units. *Brain and Language*, *115*, 189-201.

Chang, C. H. C., Lin, T. H., & Kuo, W. J. (2019). Does phonological rule of tone substitution modulate mismatch negativity? *Journal of Neurolinguistics*, *51*, 63–75.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*, *1*(5), 351–353.

Choi, W., Tong, X., Gu, F., Tong, X., & Wong, L. (2017). On the early neural perceptual integrality of tones and vowels. *Journal of Neurolinguistics*, *41*, 11-23.

Debener, S., Kranczioch, C., Herrmann, C. S., & Engel, A. K. (2002). Auditory novelty oddball allows reliable distinction of top-down and bottom-up processes of attention. *International Journal of Psychophysiology*, *22*, 309-321.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21.

Donchin, E., & Isreal, J. B. (1980). Event-related Potentials and Psychological Theory. *Progress in Brain Research*, *54*, 697–715.

Donchin, E., Ritter, W., & McCallum, W. C. (1978). Cognitive psycho-physiology: The endogenous components of the ERP. In E. Callaway, P. Tueting, & S. H. Koslow (Eds.), *Event-Related Brain Potentials in Man* (pp. 349–411). New York: Academic Press.

Elert, C. (1972). Tonality in Swedish: Rules and a list of minimal pairs. In E. S. Firschow, K. Grismtad, N. Hasselmo, & W. O'Neil (Eds.), *Studies for Einar Haugen* (pp. 151–173). The Hague: Mouton.

Engstrand, O. (1995). Phonetic Interpretation of the Word Accent Contrast in Swedish. *Phonetica*, *52*, 171–179.

Engstrand, O. (1997). Phonetic Interpretation of the Word Accent Contrast in Swedish: Evidence from Spontaneous Speech. *Phonetica*, *54*, 61–75.

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, *45*, 152–170.

Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, *50*, 259-281.

Friederici, A. D. (2017). *Language in our brain: The origins of a uniquely human capacity*. MIT Press.

Friederici, A. D., & Weissenborn, J. (2007). Mapping sentence form onto meaning: The syntax-semantic interface. *Brain Research*, *1146*, 50–58.

Gårding, E., & Lindblad, P. (1973). Constancy and variation in Swedish word accent patterns. In *Working Papers* (Vol. 7, pp. 36–110). Lund University.

Gussenhoven, C., & Bruce, G. (1999). Word prosody and intonation. In H. Van der Hulst (Ed.), *Word prosodic systems in the languages of Europe* (pp. 233–271). Berlin: Mouton de Gruyter.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language & Cognitive Processes*, *8*(4), 439–483.

Hanna, J., Cappelle, B., & Pulvermüller, F. (2017). Spread the word: MMN brain response reveals whole-form access of discontinuous particle verbs. *Brain and Language*, *175*, 86-98.

Hanna, J., Mejias, S., Schelstraete, M. A., Pulvermüller, F., Shtyrov, Y., & van der Lely, H. K. J. (2014). Early activation of Broca's area in grammar processing as revealed by the syntactic mismatch negativity and distributed source analysis. *Cognitive Neuroscience*, *5*(2), 66–76.

Hanna, J., Shtyrov, Y., Williams, J., & Pulvermüller, F. (2016). Early neurophysiological indices of second language morphosyntax learning. *Neuropsychologia*, *82*, 18–30.

Heldner, M. (2001). *Focal accent – f0 movements and beyond*. Umeå universitet.

Holcomb, P. J., & Neville, H. J. (1990). Auditory and Visual Semantic Priming in Lexical Decision: A comparison Using Event-Related Brain Potentials. *Language and Cognitive Processes*, *5*(4), 281–312.

Jasper, H. H. (1958). The 10/20 Electrode System of the International Federation. In *Electroencephalography and Clinical Neurophysiology*, *10*(2), 370–375.

Kaan, E. (2007). Event-Related Potentials and Language Processing: A Brief Overview. *Language and Linguistics Compass*, *16*(10), 571–591.

Kappenman, E. S., & Luck, S. J. (2012). ERP Components: The Ups and Downs of Brainwave Recordings. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford Handbook of Event-Related Potential Components* (pp. 3–30). New York/Oxford: Oxford University Press.

Kiesel, A., Miller, J., Jolicœur, P., & Brisson, B. (2008). Measurement of ERP latency differences : A comparison of single-participant and jackknife-based scoring methods. *Psychophysiology*, *45*, 250–274.

Korpilahti, P., Krause, C. M., Holopainen, I., & Lang, A. H. (2001). Early and late mismatch negativity elicited by words and speech-like stimuli in children. *Brain and Language*, *76*, 332–339.

Koyama, S., Gunji, A., Yabe, H., Oiwa, S., Akahane-Yamada, R., Kakigi, R., & Näätänen, R. (2000). Hemispheric lateralization in an analysis of speech sounds. *Cognitive Brain Research*, *10*, 119-124.

Kujala, A., Alho, K., Valle, S., Sivonen, P., Ilmoniemi, R. J., Alku, P., & Näätänen, R. (2002). Context modulates processing of speech sounds in the right auditory cortex of human subjects. *Neuroscience Letters*, *331*, 91-94.

Kujala, T., Tervaniemi, M., & Schröger, E. (2007). The mismatch negativity in cognitive and clinical neuroscience: Theoretical and methodological considerations. *Biological Psychology*, *74*, 1–19.

Kutas, M., & King, J. W. (1995). Who Did What and When? Using Word-and Clause-Level ERPs to Monitor Working Memory Usage in Reading. *Journal of Cognitive Neuroscience*, *7*(3), 376–395.

Kutas, M, & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203-205.

Kutas, Marta, & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470.

Lahiri, A., Wetterlin, A., & Jönsson-Steiner, E. (2005). Lexical specification of tone in North Germanic. *Nordic Journal of Linguistics*, *28.1*, 61–96.

Leminen, A., Leminen, M., Kujala, T., & Shtyrov, Y. (2013). Neural dynamics of inflectional and derivational morphology processing in the human brain. *Cortex*, *49*, 2758-2771.

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 1–14.

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique* (2nd ed.). MIT Press.

Luck, S. J., & Kappenman, E. S. (2012). *The Oxford Handbook of Event-Related Potential Components*. (S. J. Luck & E. S. Kappenman, Eds.). New York/Oxford: Oxford University Press.

MacGregor, L. J., & Shtyrov, Y. (2013). Multiple routes for compound word processing in the brain: Evidence from EEG. *Brain and Language*, *126*, 217–229.

Mäntysalo, S., & Näätänen, R. (1987). The duration of a neuronal trace of an auditory stimulus as indicated by event-related potentials. *Biological Psychology*, *24*, 183–195.

McGee, T. J., King, C., Tremblay, K., Nicol, T. G., Cunningham, J., & Kraus, N. (2001). Long-term habituation of the speech-elicited mismatch negativity. *Psychophysiology*, *38*(4), 653–658.

Menning, H., Zwitserlood, P., Schöning, S., Hihn, H., Bölte, J., Dobel, C., Mathiak, K., Lütkenhöner, B. (2005). Pre-attentive detection of syntactic and semantic errors. *NeuroReport*, *16*(1), 77–80.

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, *48*, 229–240.

Morris, J., & Holcomb, P. J. (2005). Event-related potentials to violations of inflectional verb morphology in English. *Cognitive Brain Research*, *25*, 963–981.

Myrberg, S. (2010). *The Intonational Phonology of Stockholm Swedish*. Stockholm University.

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

Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, *42*, 313–329.

Näätänen, Risto. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, *38*, 1–21.

Näätänen, Risto, & Winkler, I. (1999). The Concept of Auditory Stimulus Representation in Cognitive Neuroscience. *Psychological Bulletin*, *125*(6), 826–859.

Newman, A. J., Ullman, M. T., Pancheva, R., Waligura, D. L., & Neville, H. J. (2007). An ERP study of regular and irregular English past tense inflection. *NeuroImage*, *34*, 435–445.

Nunez, M. D., & Srinivasan, R. (2016). Electroencephalography (EEG): neurophysics, experimental methods, and signal processing. In H. Ombao, M. Linquist, W. Thompson, & J. Aston (Eds.), *Handbook of Statistical Methods for Brain Signals and Images* (pp. 175–197). CRC press.

Nunez, P. L., Nunez, M. D., & Srinivasan, R. (2019). Multi-Scale Neural Sources of EEG: Genuine, Equivalent, and Representative. A Tutorial Review. *Brain Topography*, *32*(2), 193–214.

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

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785–806.

Pataky, T. C., Robinson, M. A., & Vanrenterghem, J. (2016). Region-of-interest analyses of one-dimensional biomechanical trajectories: bridging 0D and 1D theory, augmenting statistical power. *PeerJ*, *4*, 1–12.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*, 195–203.

Pion-Tonachini, L. (2019). ICLabel tutorial: EEG Independent Component Labeling. University of California, San Diego: Swartz Center for Computational Neuroscience. Retrieved from https://labeling.ucsd.edu/tutorial/labels

Pulvermüller, F. (2001). Memory Traces for Words as Revealed by the Mismatch Negativity. *NeuroImage*, *14*, 607–616.

Pulvermüller, F., & Assadollahi, R. (2007). Grammar or Serial Order?: Discrete Combinatorial Brain Mechanisms Reflected by the Syntactic Mismatch Negativity. *Journal of Cognitive Neuroscience*, *19*(6), 971–980.

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

Pulvermüller, F., & Shtyrov, Y. (2003). Automatic processing of grammar in the human brain as revealed by the mismatch negativity. *NeuroImage*, *20*, 159–172.

Pulvermüller, F., & Shtyrov, Y. (2006). Language outside the focus of attention: The mismatch negativity as a tool for studying higher cognitive processes. *Progress in Neurobiology*, *79*, 49–71.

Pulvermüller, F., Shtyrov, Y., Ilmoniemi, R. J., & Marslen-Wilson, W. D. (2006). Tracking speech comprehension in space and time. *NeuroImage*, *31*, 1297–1305.

Riad, T. (1998). The origin of Scandinavian tone accents. *Diachronica*, *XV*(1), 63–98.

Riad, T. (2006). Scandinavian accent typology. *Language Typology and Universals*, *59*(1), 36–55.

Riad, T. (2009). The morphological status of accent 2 in North Germanic simplex forms. In *Nordic Prosody X, Helsinki* (pp. 205–216). Frankfurt am Main: Peter Lang.

Riad, T. (2012). Culminativity, stress and tone accent in Central Swedish. *Lingua*, *122*(13), 1352–1379.

Roll, M., Söderström, P., & Horne, M. (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.

Rösler, F., Pütz, P., Friederici, A., & Hahne, A. (1993). Event-related brain potentials while encountering semantic and syntactic constraint violations. *Journal of Cognitive Neuroscience*, *5*(3), 345–362.

Sams, M., Alho, K., & Näätänen, R. (1984). Short-Term Habituation and Dishabituation of the Mismatch Negativity of the ERP. *Psychophysiology*, *21*(4), 434–441.

Schremm, A., Novén, M., Horne, M., & Roll, M. (2019). Brain responses to morphologically complex verbs : An electrophysiological study of Swedish regular and irregular past tense forms. *Journal of Neurolinguistics*, *51*, 76–83.

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

Shtyrov, Y., & Pulvermüller, F. (2002). Memory traces for inflectional affixes as shown by mismatch negativity. *European Journal of Neuroscience*, *15*, 1085–1091.

Söderström, P., Horne, M., Frid, J., & Roll, M. (2016). Pre-Activation Negativity (PrAN) in Brain Potentials to Unfolding Words. *Frontiers in Human Neuroscience*, *10*, 1–11.

Söderström, P., Horne, M., Mannfolk, P., Westen, D. Van, & 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*, 119–126.

Söderström, P., Horne, M., & Roll, M. (2017). Stem Tones Pre-activate Suffixes in the Brain. *Journal of Psycholinguistic Research*, *46*(2), 271–280.

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

Sokolov, E. N. (1963). Higher nervous functions: The orienting reflex. *Annual Review of Physiology*, *1*, 545–580.

Spencer, K. M., Dien, J., & Donchin, E. (2001). Spatiotemporal analysis of the late ERP responses to deviant stimuli. *Psychophysiology*, *38*, 343–358.

Strömmer, J. M., Tarkka, I. M., & Astikainen, P. (2014). Somatosensory mismatch response in young and elderly adults. *Frontiers in Aging Neuroscience*, *6*, 1–9.

Swaab, T. Y., Ledoux, K. C., Camblin, C., & Boudewyn, M. A. (2012). Language-Related ERP Components. In E. S. Kappenman & S. J. Luck (Eds.), *The Oxford Handbook of Event-Related Potential Components* (pp. 397–440). New York/Oxford: Oxford University Press.

Teleman, U., Hellberg, S., & Andersson, E. (1999). *Svenska akademiens grammatik*. Stockholm: Svenska akademien / Norsteds ordbok.

Van Zuijen, T. L., Simoens, V. L., Paavilainen, P., Näätänen, R., & Tervaniemi, M. (2006). Implicit, intuitive, and explicit knowledge of abstract regularities in a sound sequence: An event-related brain potential study. *Journal of Cognitive Neuroscience*, *18*(8), 1292–1303.

Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience*, *170*, 223-231.

Yue, J., Bastiaanse, R., & Alter, K. (2014). Cortical plasticity induced by rapid Hebbian learning of novel tonal word-forms: Evidence from mismatch negativity. *Brain and Language*, *139*, 10-22.

Zora, H., Riad, T., Schwarz, I., & Heldner, M. (2016). Lexical Specification of Prosodic Information in Swedish : Evidence from Mismatch Negativity. *Frontiers in Neuroscience*, *10*, 1–12.

Zora, H., Riad, T., & Ylinen, S. (2019). Prosodically controlled derivations in the mental lexicon. *Journal of Neurolinguistics*, *52*, 100856.

Other sources:

Swedish PAROLE corpus, https://spraakbanken.gu.se/eng/resource/parole