Perceiving (un-)predictable words: The saliency of anomalous prosody as modulated by semantic constraints on the sentence level

An EEG/ERP study

Jule Nabrotzky

Supervisors: Dr. Mechtild Tronnier, Prof. Dr. Mikael Roll

Centre for Language and Literature, Lund University

MA in Language and Linguistics, Phonetics

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

26 May 2023

Abstract

Sentences can vary in how predictable they are, and these variations of predictiveness have been shown to influence word perception in neurolinguistic studies. However, there is ongoing debate on whether the phonological content of a word is predicted as well if the sentence context makes this possible. While some studies argue that phonological perception only happens bottom-up and is then forwarded to higher levels of analysis (e.g., to confirm or update the semantic prediction), other studies state that phonological forms themselves are predicted top-down from semantic predictions. This study tests this question on Stockholm Swedish native speakers and uses a novel approach to the problem of phonological prediction by varying both the predictiveness of the sentence, as well as the phonological realisation of the target words. In contrast to other studies, the phonological variation that is introduced to the words is a non-standard *word accent contour*, a Swedish word-level pitch accent with limited lexicality. Using EEG methodology, the study tests whether the predictability of a word modulates the amplitude of event-related potential components related to mismatch, leading to differing percepts in bottom-up vs. top-down perception. The results show that the predictive context does not significantly modulate the way a non-standard prosodic contour is perceived, however, the data shows some indication that it is the high constraint context in which it is more salient. Additionally, the type of manipulation of the prosodic contour influences the results that are obtained, indicating potential for more research into the nature of phonological word-form predictions.

Keywords: predictive coding, phonological prediction, semantic constraint, event-related potentials, Swedish word accent

Acknowledgements

While I sometimes like to pretend that I can do everything on my own, I think it's important to acknowledge that this thesis was made possible only through the help of many invaluable people. First and foremost, a big thank you to my supervisors Mechtild Tronnier and Mikael Roll for the immense support, encouragement, empathy that you've given me, as well as for the laughs and enjoyable talks that we shared. I would also like to thank Jordan Zlatev for coordinating the Master's Programme "Language and Linguistics", I think we all can be thankful to have had such an involved and caring coordinator who encourages us to see the bigger picture.

In a speech perception study, just like in real life, nothing goes without the speakers and the listeners. Therefore, thank you to Anita McAllister for agreeing to be the "voice" behind my stimuli sentences and lending your Stockholm Swedish voice for my participants to listen to. I would also like to acknowledge the support of the Division of Speech and Language Pathology at Karolinska Institutet in Stockholm for providing the recording space. Then, of course, I would like to thank all my participants for agreeing to help me with the study and being such good sports! Thank you also to the Humanities Laboratory at SOL for providing the space and materials for the EEG recordings, as well as tech and administration support.

I would also like to thank my fellow EEG-scholars Ann Hermansson, Huayuan Cui and Katja Kopaeva. Without your help with the capping, tech advice and companionship, the times in the lab would have been a lot less enjoyable. Thank you also to Louis Smit for being a good friend, fellow outdoors enthusiast, accent nerd and, last but not least, a critical opponent.

Finally, my friends and family: Thank you to Martina Belzer for being such a stellar friend and a constant in my studies, from the first day in my Bachelor's degree to now. Also, the stimuli sentences only exist because of your idea with the opposites. Thank you to Otto Maier for always encouraging me and challenging me to think differently, and to Lolina Neumeier for being a better friend than I could have ever hoped for, keeping me social and feeding me pancakes on the weekends. Und zu guter Letzt, danke Mama, Papa, Hannah und Mia, dass ihr jeden meiner Umzüge unterstützt habt und immer an mich geglaubt habt. Bald bin ich wieder zuhause.

Table of contents

List of figures

List of tables

Abbreviations

EEG Electroencephalography

ERP Event-related potential

LMM Linear mixed models

MMN Mismatch negativity

N400 A negative-going ERP components at 400ms that responds to semantic violations

PMN Phonological Mismatch Negativity

PrAN Pre-Activation Negativity

P600 A positive-going ERP component at around 600ms that is related to syntactic violations

Chapter 1 Introduction

The most fundamental task that listeners of spoken language need to accomplish is segmenting the incoming acoustic signal of speech into chunks and units that form words. However, contrary to how it is often perceived, the signal does not clearly delineate these segments (Port & Leary, 2005), regularly omits segments (Ernestus et al., 2002) and delivers a wealth of additional information, such as speaker characteristics, voice quality and intonation (Garellek, 2019). This fundamental task therefore poses quite a challenge, and many mechanisms have been proposed that help a listener deal with the contextual variability of speech signals so that the segmental and suprasegmental information can be integrated in the larger context (Brouwer et al., 2013, p. 520).

For example, connectionist accounts, like the TRACE (McClelland & Elman, 1986) or Shortlist (Norris, 1994) model propose that each word form is stored only once in the mental lexicon and that pre-lexical processes help normalize the input to the mental lexicon (Mitterer & Blomert, 2003). On the other hand, episodic models of speech perception (Goldinger, 1998; Johnson, 1997) postulate that multiple exemplars of words are stored in the mental lexicon in forms that are reduced to different degrees. In addition to disagreements on the nature of word storage in the mental lexicon, these models also disagree on whether information can flow top-down from higher levels of representation down to the level of phonemes (McClelland & Elman, 1986) or if the information flow can happen in a purely bottom-up fashion (Marslen-Wilson, 1987).

Essentially, top-down processing entails that predictions about the upcoming linguistic content are taken into account when processing lower-level input. In recent years, evidence for top-down processing has accumulated, with both behavioural and neurological studies arguing that their results show differences in processing of predictable vs. unpredictable stimuli (Kuperberg & Jaeger, 2016). Due to these promising results, the theory of predictive coding has gained traction in the field (Bastos et al., 2012; Friston & Kiebel, 2009). Predictive coding, like the TRACE model (McClelland & Elman, 1986), is based on Helmholtz’s (1896) concept of unconscious inference and states that perception is based on mental models called hierarchical generative models. These pass on top-down-information from the predictive mechanisms toward the sensory organs, anticipating a certain sensory experience and only forwarding bottom-up feedback when the sensory input deviates from the representation on this lower level generated by the prediction, a *prediction error* (Bastos et al., 2012, p. 704).

While several studies have utilized predictive coding theory to explain speech perception, the evidence is not yet conclusive. On the one hand, phenomena like phonemic restoration (Petkov & Sutter, 2011; Samuel, 1996), where a segment or tone masked by a cough is recovered in the percept so that the listener reports having heard the masked segment, show that representational units can be the basis of perception in the absence of the expected bottom-up input. However, this phenomenon disappears when the distorted segment is too large to be explained away by predictions (Riecke et al., 2009). Additionally, it seems that this process is modulated by sentence context and attentive processes (Brouwer et al., 2013; Riecke et al., 2009).

When investigating larger language units, the interpretation of effects gets increasingly difficult. Predictive processing has in fact been investigated for words and sentences in great detail over the last years, especially using the electroencephalographic/event-related potential (EEG/ERP) methodology (Kutas & Federmeier, 2011). The N400, a negative-going voltage at around 400 ms past word onset, has been especially well linked to the prediction of words. However, even though it has been thirty years (Connolly & Phillips, 1994) since experiments were devised to test whether or not word-form predictions are generated by the predictive mechanisms of the brain (which would be consistent with predictive coding), the evidence is still unconclusive (Nieuwland et al., 2018).

Specifically, it is difficult to disentangle whether the effects that are observed are due to difficulties in word recognition or integration of the new information as opposed to predictive processing-related error signals (Van Petten et al., 1999). To test the theory of predictive coding for language processing further, these possibilities need to be disentangled. Therefore, this study aims to provide evidence for predictive coding in sentence comprehension by bridging the gap of evidence between studies of lower-level tonal processing and higher-level sentence processing. To achieve this, a study design that was originally used to test mismatching allophones

(Brunellière & Soto-Faraco, 2015) is adapted. However, instead of phonemes, it uses less categorical prosodic contours, more specifically, Swedish word accent contours.

These are prosodic tonal contours that are assigned to all Swedish words based on morphological factors (Riad, 2013, p. 192). They are well suited to exploring predictive processing since their validity as a cue for word recognition is limited due to the small amount of Swedish minimal pairs that only differ between Accent 1 and Accent 2 (Elert, 1972). Still, they have been shown to facilitate word processing by eliminating possible continuations for a word beginning (Roll, 2022), showing that they are relevant to speech perception as a whole. Additionally, their realization is affected by whether or not they receive focus on the sentence level, meaning that they are depenent on sentence-level processes (Bruce, 1977; Myrberg & Riad, 2015).

The study explores the perception of Swedish words in two sentence contexts, either a highly predictive semantic context which limits the sentence-final word to only a few viable scenarios, or an unpredictive sentence context that has a high number of possible sentence completions. In both cases, the final word will be modified in half of the cases to contain an anomalous prosodic contour. The thesis thereby aims to contribute to our understanding of phonological predictions as modulated by semantic context constraints by asking the following research question:

> *How does the predictive power of the semantic context in which a word is embedded modulate the way anomalous melodic contours in that word are perceived?*

Predictive coding theory implies that there is a difference between the perception of anomalous prosodic contours in lowly and highly constraining contexts. However, as discussed with regards to the phonemic restoration effect, it seems to depend on the magnitude of the anomaly whether it is restored or it leads to a prediction error. Based on this knowledge, the following hypotheses are generated for the research:

> H0: The sentence context in which the target word is embedded makes no difference in the perception of typical vs. anomalous prosodic contours.
>
> H1: Anomalous prosodic contours are more perceptually salient in lowly constraining than in highly constraining sentence contexts.

H2: Anomalous prosodic contours are more perceptually salient in highly constraining sentence contexts.

Before the study is described in more detail, chapter two of this thesis first gives a more detailed background on predictive coding and the experimental approaches that have been used to examine its validity. Since the thesis builds on the prosodic expectations of Swedish speakers, the chapter then illustrates Swedish prosody in more detail and then takes a closer look at the research that has already been done in the field of phonological prediction in event-related potential studies. Chapter three then explains the precise methodological approach that is taken to answer the research questions that were posed above. The results of the experiment are presented in chapter four and discussed in chapter five, together with a reflection on the limitations of the study and possibilities for further research. To conclude, chapter six relates the results back to the research questions and hypotheses and reflects on the relevance of the study for the larger research context.

Chapter 2 Literature review

This chapter reviews some of the research that has previously been conducted in the field of phonological prediction and predictive coding and introduces Swedish word accents. It first takes a closer look at predictive coding theory and how it can be applied to speech perception. The chapter especially focuses on the differing ways predictive coding has been interpreted in a language context. The chapter then describes the phonology of Swedish word accents to give background on the way the stimulus sentences in the study were constructed. The third subchapter reviews the efforts to understand phonological prediction and how it can be observed with an EEG/ERP approach.

2.1 Predictive coding

2.1.1 General concepts

Predictive coding is a computational framework that aims to explain perceptual processes in the human brain. However, it originated as a framework for the efficient encoding of information in computational processing (Clark, 2013, pp. 182–183). For example, when transferring an image from one unit to the other, it is uneconomical to transmit information about the colour of every single pixel. Instead, if it is assumed that each pixel is surrounded by pixels of the same colour, information about a pixel only needs to be encoded when it violates this assumption. Through this, the processing cost of transferring an image is greatly reduced.

When applied to mechanisms in the brain, predictive coding works in much the same way. In its basic form, it posits that information processing aims to minimize the amount of bottom-up input that needs to be processed. The system achieves this by building a model that aims to best explain the sensory input, minimizing prediction error (Lupyan & Clark, 2015, p. 279). The system is updated with the help of prediction errors to arrive at the best model using Bayesian inference (Brown & Friston, 2012). Mathematically, this is achieved by minimizing surprise, meaning that the system tries to incorporate knowledge about the plausibility of a certain sensory input on the basis of available information (Brown & Friston, 2012, p. 3).

Readily perceivable evidence for predictive coding in the visual modality comes from phenomena like the flash-lag effect: A flashing object is perceived as lagging behind a constantly visible object when both are moving at the same speed (Khoei et al., 2017). Predictive coding theory explains that this is because our perception of the constantly visible object is not based on bottom-up visual data, but our prediction of where the object will be is extrapolated from our knowledge of its current speed. The presence of a flashing object has higher predictive uncertainty, leading to a percept that relies more on bottom-up cues, which take longer to process, leading to the perceived lag (Khoei et al., 2017).

In its basic tenets, predictive coding is a mechanistic theory about cortical processing (Kogo & Trengove, 2015). The model argues for the functional separation of error neurons (the superficial pyramidical cells) and neurons that represent the sensory input prediction (deep pyramidal cells) (Clark, 2013, p. 188; Friston & Kiebel, 2009). However, as also described in the introduction, predictive coding runs into a problem when articulating test hypotheses for experiments since it is often not clear what the neural activity observed for unexpected stimuli represents. On the one hand, it could be generated by so-called “error neurons” that propagate the information about the unexpected sensory input to higher levels. However, because predictive coding is thought to be iterative, the system will update according to the error signals to arrive at a model that better explains the sensory input (Kogo & Trengove, 2015). It could therefore also be the case that the neural activity that is measured is the result of updating the representations of the model at lower levels. Therefore, evidence for predictive coding requires careful experimental designs with neuroimaging methods, one of which is electroencephalography.

2.1.2 Testing the theory of predictive coding with event-related potentials

Understanding the processing of sensory input through electroencephalography (EEG) has been a fruitful method for neurolinguistic research mainly because of its high temporal resolution. In the EEG methodology, electrodes are placed on the scalp of a subject and voltage changes at the electrodes are measured over time. This can be done at a much higher sampling frequency than, for example, functional magnetic resonance imaging (fMRI), providing detailed information on the time course of neurological processes. However, this approach has two main shortcomings: Since the electrical activity at the scalp is measured, only a certain type of brain activity can even be detected by non-invasive EEG. The main source of the EEG voltage changes comes from post-synaptic potentials of cortical pyramidal cells (Luck, 2014, pp. 40–41). Additionally, since voltage is, by definition, relative to the ground and reference, and voltage changes can be influenced by one or by several generators of electrical activity, the EEG methodology is not particularly well suited for source localisation (Luck, 2014, pp. 48–52). Therefore, EEG signals are mainly described using ERP "components", positive- or negative going components that occur a certain amount of time after an event that is relevant to the research question.

Evidence for predictive coding in language using the EEG/ERP methodology primarily comes from studies on pitch processing (Baldeweg, 2006; Kumar et al., 2011; Wacongne et al., 2011). For example, when presenting a constantly appearing stimulus tone followed by a different tone (an MMN paradigm, see subchapter 2.3.1 for more details), Haenschel et al. (2005) found that the number of standard stimuli preceding the different tone modulated an early slow positive wave in the ERP signal. This effect was interpreted as the neural correlate of memory trace formation and appears similar to effects found for the visual system (Baldeweg, 2006). In other words, the more often the stimulus tone is repeated, the more the sensory system is anticipating a similar stimulus, shown by adaptations in very early ERP components.

Wacogne et al. (2011) also adapted the MMN paradigm by showing that errors are hierarchical and can be layered. When they presented a series of tones with a final mismatching tone several times and then omitted the mismatching tone at the end or replaced it with a matching tone, they observed two kinds of error signals on top of each other. On the one hand, the mismatching tone itself generated an expectation of the mismatching tone, on the other hand, that expectation was violated by the silence. The results showed that the MMN at the final mismatching tone was modulated by how often the original tone series was played beforehand and showed that, when presented often enough with the original sequence, the replacement of the mismatching tone with a matching tone resulted in an MMN. But most importantly, the omission of the final mismatching tone created a bigger MMN, which the authors propose represent the layering of the expectation of the mismatching tone with the error signal of the omission.

However, as mentioned in the introduction, extrapolating the model to account for higher-level language-related phenomena suffers from certain pitfalls that have also been found for general higher-order cognitive processes. Because of this, it is still a challenge to generalize effects observed in basic mismatch paradigms to higher-level processes that are modulated by more abstract types of experience (Firestone & Scholl, 2016). The study presented in this thesis approaches this problem by utilizing a linguistic unit that is more similar to the tones used in the mismatch paradigms, namely Swedish prosodic contours on the word level, also called *word accents*.

2.2 Swedish word accents

In this study, Swedish word accent contours appear in varying forms on the target word to create a percept that is anomalous. Before this literature review gives more details about studies that have been conducted to test language-related prediction, this subchapter is meant as an introduction to the linguistic phenomenon in the study so that the prior studies can be understood in relation to it. Additionally, this subchapter is meant as a guide to better understand the manipulations of the word accent and their effect, as well as the function of word accents in Swedish speech perception. To achieve this, the subchapter gives a short background on word accents in the Swedish language, along with some comments on sentence-level prosody as it pertains to the stimuli sentences that were used in the experiment.

2.2.1 Typology of Swedish word accents

Swedish is a language with lexical stress that also has tonal contours assigned to every word, called "word accents" in the Scandinavian tradition (Bruce, 2007, p. 116). In cross-linguistic typologies, these accents are also often called *pitch accents* and distinguished from stress-accent and tone (Jun, 2005), although it is not entirely clear how to delineate them from tones[1] like those that appear in, e.g., Mandarin Chinese (Hyman, 2009). There are two basic tonal contours, called Accent 1 and Accent 2 or acute and grave, whose F0 contour differs between regional varieties of Swedish (Gårding & Lindblad, 1973). Some varieties of Swedish, like Finland Swedish, do not have this tonal distinction at all (Bruce, 2007, p. 117; Gårding & Lindblad, 1973). Additionally, there is a focal distinction in the word accents, meaning that the tonal contour of the word accent is different depending on whether or not the word is focally accented on the sentence level (Bruce, 1977).

For Stockholm Swedish, the focally accented version of the word accent is two-peaked, meaning a second F0 peak is added after the first F0 peak. The difference between word Accents 1 and 2 in Swedish is the timing of the F0 peaks in relation to the stressed syllable of the word. In Accent 1, the second peak is aligned with the stressed syllable. The focal version of these words (where the second peak is present) therefore generally shows a rising contour through the stressed syllable. The first peak is located in the pre-stress syllable, although this peak is sometimes elided. This means that Accent 1 words without a focal accent are characterized by an accentual fall on the stressed syllable. In Accent 2 words, the second peak is aligned with the non-stressed syllable (or secondary stressed syllable in compounds). The first F0 peak is aligned with the stressed syllable and occurs slightly earlier than the second peak in Accent 2 words, leading to a high falling contour in the stressed syllable (Bruce, 1977, 2007; Myrberg & Riad, 2015, p. 116).

---

[1] Just like the word "accent", the word "tone" has a general and a more linguistic meaning. In general, the thesis has used the word tone in its general meaning (a sound with a certain pitch, like a A_4 in music (440Hz). On this occasion it is used in its linguistic meaning. Any latter mentions of tone again refer to the general meaning.

Table 1 shows the tonal contours of the four types of word accent according to autosegmental notation with the addition of a superscript *F* to denote that the peak is a focal peak.

Table 1: Swedish word accent contours

Swedish word accents in Stockholm Swedish adapted from (Myrberg & Riad, 2015, p. 116)			
Sentence context	Accent 1	Accent 2	Accent 2 (compound)
focal	L*HF	H*LHF	H*LHF*
non-focal	HL*	H*L	H*L

On an articulatory level, recent evidence suggests that while Accent 1 is aligned with the lip aperture, Accent 2 is more aligned with the tongue body movement, although this alignment is less stable (Svensson Lundmark et al., 2015). There are also many acoustic measurements that differ between Accent 1 and 2 words, including durational patterns in the stressed syllable (Heldner, 2001; Svensson Lundmark et al., 2017) and more creaky voice in Accent 1 words (Svensson Lundmark et al., 2017).

There are some regularities to explain which word accent is assigned to which word, with both morphology and post-lexical prosody playing a role. Firstly, all monosyllabic words are assigned Accent 1. However, these monosyllabic words or stems can be affixed with suffixes, some of which carry Accent 2, which is then realized on the stem of the suffixed word. Some suffixes that carry Accent 2 are the infinitive, plural and past tense suffixes, as well as some present tense forms (-ar). That means that the verb *känna* (eng. "to feel") carries a different word accent depending on its conjugation as seen in example (1). Additionally, adjectives and nouns can carry different word accents depending on number as seen in (2). Compounds with two stressed syllables also receive Accent 2, although in this case the focal peak is situated on the secondary stressed syllable as seen in (3).

(1)	^{2}*kän-na*	^{1}*kän-ner*	^{2}*kän-de*	^{1}*kän-t*
	feel-INF	feel-PRS	feel-PST	feel-PTCP
	‘to feel’	‘feels’	‘felt’	‘felt’

(2)	^{1}*enkel*	^{2}*enk-la*	^{1}*häst-en*	^{2}*häst-ar*
	easy	easy-PL	horse-DEF	horse-PL
	‘easy’	‘easy’	‘the horse’	‘horses’

(3)	^{2} *ˈsommar-terˌmin-en*
	summer-term-DEF
	‚the summer term‘

2.2.2 The function of Swedish word accents

Since the word accents do not carry a high functional load in Swedish phonology, meaning that they do not have a distinctive function excepting a low three-digit number of minimal pairs (Elert, 1972), many of which are based on archaic or very low frequency words, their function has been called into question (Roll, 2022). Prior studies have shown that they do not significantly affect the recognition of words (Abelin & Thorén, 2015) when appearing in isolation, although a mismatching word accent contour on minimal pairs in primed contexts has been shown to raise reaction times (Althaus et al., 2021) and lead to an N400 effect (Kwon, 2023). These results indicate that while they might not carry a distinctive function, listeners do seem to glean some lexical information from them.

However, a much stronger case can be made for the predictive function of Swedish word accents regarding morphology. While not minimal pairs as such, word pairs such as ^{1}*hästen-*^{2}*hästar* or ^{1}*känna-*^{2}*känner* do sound similar and only differ in the last syllable (Riad, 2013, pp. 182–3).

Since the word accent that the stem of a inflected word carries is generally determined by the inflection, it could serve as a way to pre-activate potential word endings during the perception of a word (Roll, 2022, p. 4). The word accent could therefore carry a processing benefit by eliminating incompatible suffixes beforehand.

The evidence for this comes from several psycholinguistic studies that have mainly focused on the impact of a word accent mismatch on processing: Both in isolated words and in the sentence context, a word accent contour on the stem that mismatches the (expected) suffix leads to processing difficulties. In an fMRI study on singular and plural nouns (see example (2)) where the word accent pattern on the stem mismatched the suffix that followed, subjects showed more activation in the left inferior parietal lobe (Roll et al., 2015). Since number processing has been linked to this area, the incorrect cues from the stems might have led to difficulties in integrating the number information. Additionally, these invalidly cued suffixes lead to an increased P600 effect in event-related potential studies (Roll, 2015; Roll et al., 2010, 2013; Söderström et al., 2017).

Additionally, differing pre-activation negativity (PrAN) effects for Accent 1 and Accent 2 cues show that the two word accents have differing predictive weights. In general, word stems with Accent 1 can eliminate more potential words than words with Accent 2 (Felder et al., 2009, p. 1892; Roll, 2015, p. 149) due to the fact that more suffixes are associated with Accent 2 and Accent 2 is also the default for compound words. This means that hearing an Accent 1 stem leads to greater predictive certainty, decreasing processing times (Roll, 2015; Söderström et al., 2017).

2.2.3 Swedish word accents in relation to sentence-level prosody

The word accents in this study do not appear in isolation but are embedded into sentences. To motivate the way these sentences were constructed, this subchapter gives some background on how the word accents interact with their sentence context. As already mentioned, whether a word appears with a focal accent depends on sentence context. However, even if two words are focally accented, their actual tonal realisation of this focal accent may differ according to their position in the sentence. For example, downstepping throughout an intonational phrase, the gradual lowering of the f0 contour with every word, might affect the amplitude of the pitch movements on the focally accented word. The positioning of the word within the sentence, as well as the words surrounding it also have an effect on the F0 contour. Figure 1 illustrates this on three occasions of the word *kvinnan* (en. the woman) in the stimulus material for the study. The contour on the stressed syllable varies from a fall to a slight rise to a fall, and the contour in the unstressed syllable is either rising or level.

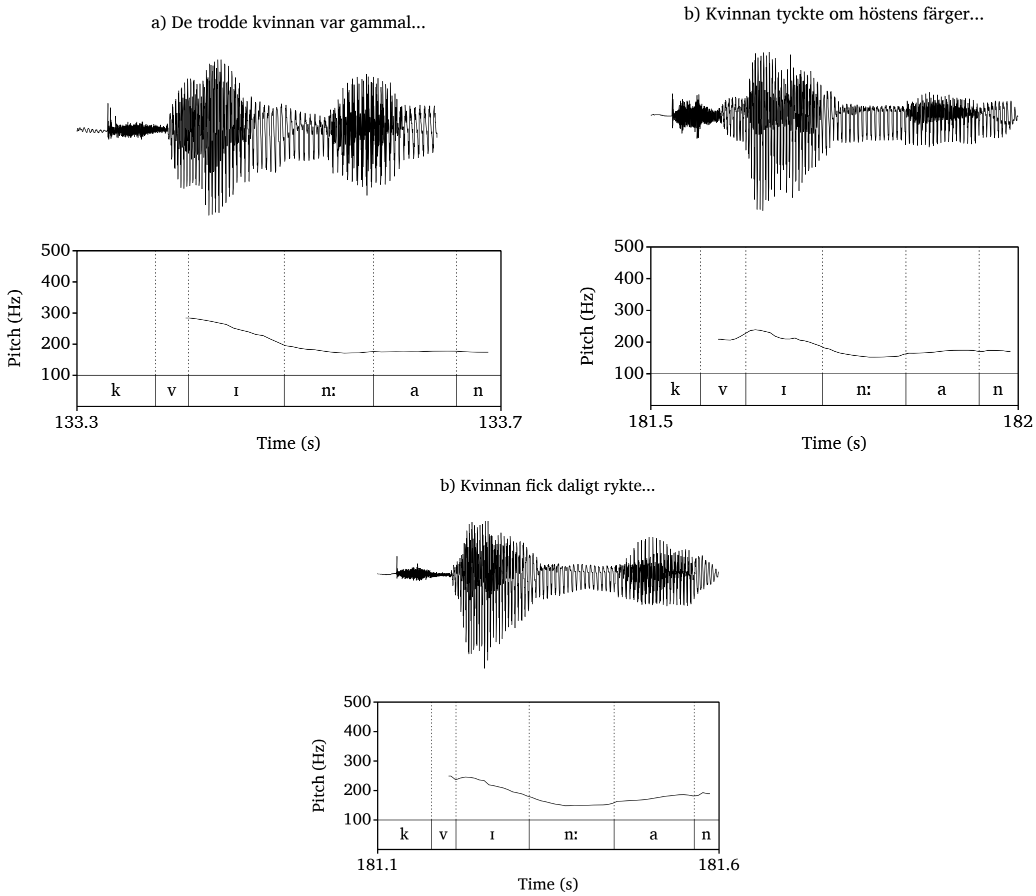


Figure 1: Three examples of the word "kvinnan" from the materials

Furthermore, the accents of adjacent words can interact with each other to form plateaus (Bruce, 1987; Myrberg, 2010, p. 25). Figure 2 shows such a plateau between focal accents in the phrase *nyfikna på butikerna* (en. curious about the shops). The word *nyfikna* carries the focally accented version of Accent 2. The first F0 peak is located on the first syllable, while the second F0 rise starts in the second syllable. However, the three syllables after this also carry a high tone before the fall in the stressed syllable /ti/ of the word *butikerna.* What is notable here is that because of the plateau, the f0 pattern of the word *butikerna* is affected. While it is stressed on the second syllable (which is also indicated by the long vowel), the first syllable contains a high tone.

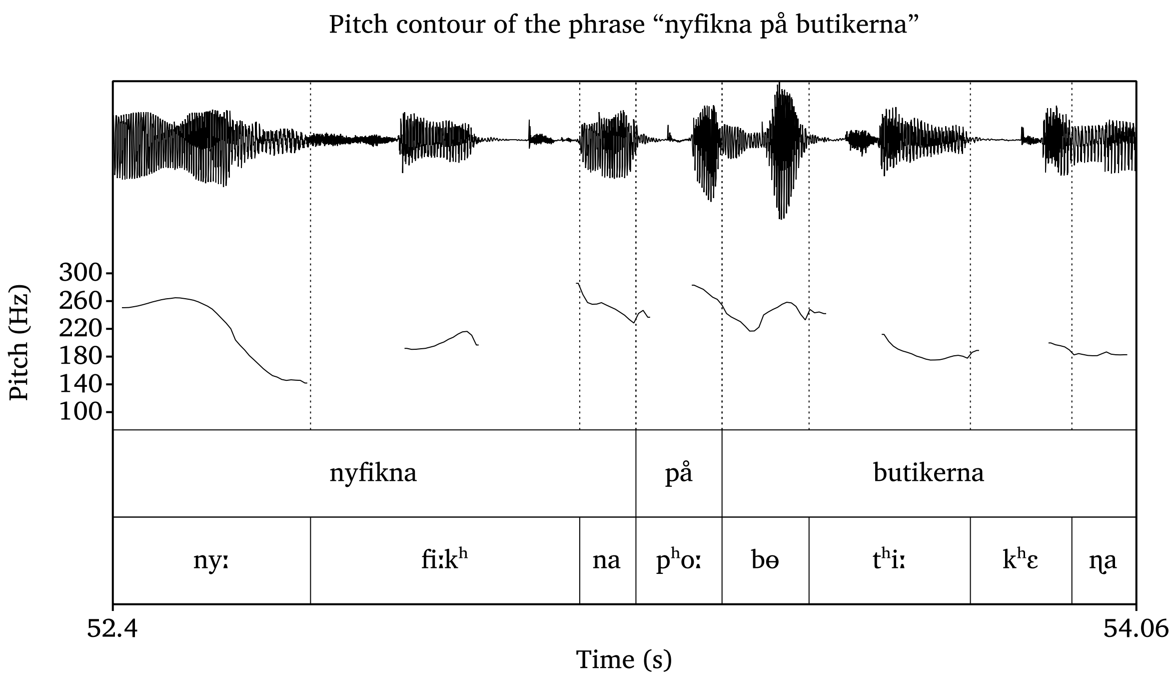


Figure 2: Example for a plateau when pitch accents are concatenated

Additionally, the information structure of a sentence also has an influence on the realization of a focally accented word. Traditionally, a focal accent is expected for pragmatic uses of focus (Krifka, 2008, p. 250), such as correcting and confirming information, highlighting parallels and providing answers to Wh-questions. However, Ambrazaitis (2009, p. 153) found that the pattern of a focal accent can be neutralized by a confirmation of previously known information. In these cases, the focal accent is characterized by a falling tone on the stressed syllable.

2.3 Measuring prediction with event-related potentials

In the study at hand, ERPs are utilised to understand how a phonological prediction caused by semantic context affects perception. However, there are many different kinds of components that have been elicited in experimental paradigms which try to measure prediction (Nieuwland, 2019). Not all of them are relevant for the experiment at hand since many of them were elicited in response to visual stimuli. This subchapter introduces the most applicable auditory ones and how they have been observed. Besides the N400 component, which has already been briefly introduced above, there are several earlier brain responses also associated with auditory stimuli, for example the Mismatch Negativity and the Phonological Mismatch Negativity.

2.3.1 The N200 component group

The simplest account of how the prediction of auditory stimuli affects the EEG signal comes from the *Mismatch Negativity (MMN)* (Näätänen et al., 1993). It is most often observed in oddball experiments that present a devious stimulus amid several homogenous stimuli (usually simple tones) and peaks at around 160-220ms. The MMN is elicited even when the auditory stimuli are task-irrelevant (Luck, 2014, p. 85). It has been source-localised as originating from the auditory cortex and the frontal cortex (Näätänen et al., 2012).

While the oddball paradigm is not very similar to natural sentence comprehension, there have been linguistic studies that elicited an MMN-like effect in relation to degraded or missing phonetic content. However, it is not clear how this MMN interacts with higher-level linguistic prediction. In an experiment where some target words were incomplete, Bendixen et al. (2014) manipulated the predictability of the sentence context that the target words were placed in. They showed a larger omission response for predicted than unpredicted words which occurred at around 150ms after word onset and was source-localised to auditory cortical areas.

However, not all degradations of phonetic content lead to such results. For example, Boulenger et al. (2011) varied the *cloze probability* of sentence completions (how often a certain word was inserted by participants in a sentence completion task) and observed its effect on the perception of normal and time-reversed target words. They found a fronto-temporal negativity in the MMN window for the time-reversed words irrespective of sentence contexts. The sentence context only affected later time-windows, where the N400 effect for time-reversed words was modulated by the cloze probability. For high cloze words, the study found bigger N400 effects for time-reversed words. However, they observed smaller negativities for time-reversed than normal words when the cloze probability was low. This was interpreted by the authors as showing that the incongruency of the target word with the expected sentence completion was less prominent when the phonetic disruption was present.

Such diverging results seem at first to be at-odds with a unified theory of prediction in language comprehension. In some cases, a violation of the prediction leads to an error signal, like with Bendixen et al.'s (2014) missing syllable codas. In other cases, the violation is not even perceived, like in phonemic restoration experiments. However, predictive coding is actually particularly well-suited to explain the processing of ambiguous and imprecise stimuli (Brown & Friston, 2012). This is because predictive coding states that the prediction errors are precision weighted. This means that depending on the expected precision of the error, some errors do not lead to an update of the model but instead are supressed. Regarding language, it has been theorized that precision-weighting can be seen by the predictive weight that linguistic cues are given in different contexts (Bornkessel-Schlesewsky & Schlesewsky, 2019, p. 8). The authors theorize that error signals pertaining a bottom-up cue are afforded a higher precision weighing if the cue is more reliable and relevant for sentence comprehension, as well as that the precision weighing of a cue can also change with the context in that it is perceived (Bornkessel-Schlesewsky & Schlesewsky, 2019, p. 9).

2.3.2 The N400 component

In linguistic studies, prediction has been most clearly demonstrated in EEG/ERP studies on the word level, where the "N400", a negative-going component at around 400 ms after a relevant word, has been shown to correspond to the predictiveness of the word (Kutas & Federmeier, 2011). More specifically, the N400 component is reduced for words that are expected, while it is large for unexpected words in a highly constraining sentence context or words in a lowly constraining sentence context. It seems that the N400 reduction represents the processing benefit that prediction can give. Phonological prediction is less well understood, and researchers interested in this face the additional challenge that these predictions are to some extent dependent on word level predictions. That is, to generate a prediction of the phonological content of the word, it is reasonable to assume that we don't just predict the next phoneme with our knowledge about what phonemes are likely to follow each other, although that certainly does play a role. Instead, we consider our knowledge about the semantic context. On the other hand, a phonological prediction is crucial for the application of predictive coding to speech perception, since it is much closer to the sensory level that generates a prediction error than the conceptual level of a word.

2.3.3 Disentangling the PMN and the N400 – early approaches

The basis for the research on phonological word form prediction in connection with the EEG/ERP methodology was laid by Connolly and Phillips (1994). In an earlier study (Connolly et al., 1992), observed that the cloze probability of a target word had an effect on both the N400 amplitude, but also an earlier negative-going component at around 200ms. However, masking the target word delayed the N400 but not the earlier component, which lead to the proposal that they are functionally separate.

To explore this further, they developed an experiment that created highly constraining sentences in four different conditions, with either a phonological mismatch, a semantic mismatch, neither mismatch (the original sentence) or both a phonological and a semantic mismatch. The semantic mismatch was created by ending the sentence with a word that started with the same sound as the highest cloze probability word but did not match semantically. The phonological mismatch was created by choosing a different word than the highest cloze word that still fit the semantics of the sentence. In the condition with both semantic and phonological mismatch, the target word was semantically anomalous and started with a different phoneme. This paradigm showed separate components for the phonological and semantic mismatches, with both components present in the double mismatch condition (Connolly & Phillips, 1994, pp. 257–8). The earlier component that was elicited by a phonological mismatch peaked negatively at around 270-300 ms. This was the first dissociation of the so-called Phonological Mismatch Negativity (PMN) from the more semantic N400.

The PMN effect was replicated in several studies (Connolly et al., 1995, 2001; Dehaene-Lambertz et al., 2000; Hagoort & Brown, 2000; van den Brink et al., 2001; van den Brink & Hagoort, 2004), for an overview, Lewendon et al. (2020). However, the issue of functional separation proves to be challenging, since some researchers argued for early semantic integration processes at the root of the PMN effect (Hagoort & Brown, 2000; van den Brink et al., 2001). Central to this debate is the timing of the components that are elicited by phonological mismatches.

For example, D'Arcy et al. (2000) used a similar paradigm to replicate the results of Connolly and Phillips' (1994) experiment, however, they used a Computerized Token Test (CTT) instead of high cloze probability sentences to elicit strong phonological predictions. They examined the N2b component in two separate time frames and found that the condition with phonological mismatch showed a more negative N2b component as the semantic mismatch condition in the earlier time frame (130-230 ms). However, the effect was reversed for the later time window (250-350 ms) (D'Arcy et al., 2000, p. 42). The authors interpret this finding as showing that the initial matching phoneme in the semantic mismatch condition delays the detection of a semantic mismatch, indicating that individual phonemes are predicted instead of the whole word. The effect of phonological similarity in semantically incongruous words was investigated in a similar study by Desroches et al. (2009) who added a rhyme condition to Connolly and Phillips' (1994) paradigm. Contrary to other studies, they used pictures of the target words to elicit phonological predictions. Their results are consistent with a separate processes-account of the N400 and PMN.

However, in a similar study with high cloze-probability sentences that examined the onset of the N400 response for semantically anomalous words in relation to the isolation points of the words, van Petten et al. (1999) found that the N400 response started before the isolation points, demonstrating that the semantic utilisation of phonological input starts immediately. The onset of an N400 response for semantically anomalous but phonologically congruent word onsets was delayed until the word diverged from the highest cloze probability word. Crucially, this study only showed a difference in N400 onset as opposed to a functionally separate PMN component.

Theoretically, the two accounts share the notion that phonemic information is analysed but differ in their interpretation of whether the incoming phonemic information was predicted and therefore elicits a negative component when the prediction isn't confirmed. The semantic account posits that early negativities simply arise from the onset of semantic processes on the basis of incoming sensory information (van den Brink et al., 2001, pp. 968–9; Van Petten et al., 1999, pp. 411–412). It is clear that the original paradigm used in Connolly and Phillips' (1994) study is insufficient to clear up the matter, which is why alternative approaches to the question have been developed.

2.3.4 New evidence and state of the field

For the purpose of disentangling phonemic prediction from semantic integration, Newman et al. (2003, p. 642) developed the "clap without the /k/" paradigm. Participants were instructed to imagine the word as described and then heard a congruent ("lap") or incongruent ("ap") word. The results showed a PMN in the incongruent conditions (Newman et al., 2003, p. 646), arguing against early-starting semantic processes as in van Petten et al. (1999). However, the authors could not exclude the possibility that an early onset P300 effect was responsible for the absence of the PMN in the congruent condition (Lewendon et al., 2020, p. 2; Newman et al., 2003, p. 646). Additionally, this also could not rule out whether the early negativity resulted from comparing the auditory input with preselected lexical candidates, as argued by Hagoort & Brown (2000). To clarify this, a follow-up study where participants deleted consonants from non-words as well as words was conducted (Newman & Connolly, 2009). Here, the N400 amplitude was modulated by the lexicality of the stimulus, adding more evidence for the functional separation-argument (Newman & Connolly, 2009, p. 119).

Another approach uses a phonological rule of English to study whether word forms are predicted based on sentence context. DeLong et al. (2005) presented listeners with sentences that cued a certain word at the end of a sentence that was preceded by an indefinite article (a/an). Since the choice of indefinite article depends on the phonological onset of the next word (a kite/an airplane), they theorized that they would see an N400 effect for articles that mismatched the predicted target word's phonological realization. Indeed, they found a negative deflection for the mismatching indefinite articles at around 200-500ms after the word onset. However, several replication studies failed to observe the same effect (Ito et al., 2017; Nicenboim et al., 2020; Nieuwland et al., 2018), citing low statistical power and the usage of the average reference as possible reasons for the effect observed in the original study (Nicenboim et al., 2020, p. 3).

In fact, methodological criticisms against many of the earlier PMN experiments have been raised, for example considering the low sample size of some of the earlier studies and the reliance on visual inspection to find time windows or electrodes that correspond to the hypothesis (Nieuwland, 2019; Poulton & Nieuwland, 2022). In their study with a relatively high sample size of n=48, Poulton and Nieuwland (2022) failed to replicate the PMN effect using the same paradigm as in Connolly and Phillips (1994). Using significantly more electrodes than in this study, they also failed to replicate a different scalp distribution in the N400 and PMN. These results call into question the already unclear evidence for the functional separation of the PMN effect. Instead, Poulton and Nieuwland (2022, p. 572) argue for a "multiple-processes-acount" of the N400, theorizing that the N400 does not just represent prediction on the semantic level, but takes other levels of processing into account. The early negativities in the PMN studies could then be understood as an early-onset N400, owing to the similar scalp distribution in the replication study.

While there are many studies on the effect of an unexpected word and how this is modulated by predictive context, there is less work on the perception of words that are expected, but that are phonetically deviant. An exception is a study on the perception of dialectal variation in lowly and highly constraining sentence contexts (Brunellière & Soto-Faraco, 2015). In their study, the authors varied the predictive value of sentence contexts that a word with either a matching regional dialect[2] or a mismatching regional dialect appeared in. The two dialects differed in their realisation of the vowel phonemes /e/ or /u/. The authors found an early effect of the phonological mismatch in the lowly constraining sentences, but not the highly constraining sentences, possibly indicating that the mismatch could be resolved faster in a highly constraining context (Brunellière & Soto-Faraco, 2015, p. 56). However, the authors could not eliminate the possibility that this effect was simply due to a delayed recognition of the target word due to the unfamiliar phoneme. In fact, recognition effects are one of the major pitfalls of experiments that try to demonstrate predictive coding for higher-level processing (Firestone & Scholl, 2016, pp. 15–16). Therefore, it is not clear if the sentence context or simply the earlier recognition led to a differing percept.

To conclude, while prediction is readily observable in paradigms that elicit the N200 or N400 components, finding a negativity related to word form predictions is less straight-forward. It seems that the effects are quite small and not very robust (Nicenboim et al., 2020), possibly requiring more advanced statistical approaches like Bayesian statistics across several studies.

2.3.5 Reactions to anomalous pitch contours

When adapting the PMN study paradigms for prosodic violations, it is useful to survey which ERP responses for such violations have been obtained in previous studies. In contrast to words or even phonemes, the F0 contour is a very straight-forward acoustic trait that has equivalents in non-language related settings (such as the tones used for MMN paradigms). However, MMN paradigms usually do not utilize "context" in the same way as it is proposed for this study, which is why an analogous study that created pitch-related expectations based on familiar melodies in music (Schön et al., 2004) might be more fitting. In fact, the study found an early anterior

---

[2] Note that the authors use the word "accent" to describe the two varieties of Spanish that are used in the study (Western and Eastern Catalan), however, since that word is used to describe prosodic contours in Swedish in this study and the literature on Swedish prosody, the word "dialect" is chosen to refer to regional variations of a standardized language to avoid confusion.

negativity peaking at around 150ms for pitch violations in musical sequences. In the same study, a similar early negativity was also evoked for pitch violations in language, created by an unusual right-edge phrase boundary tone.

When adding unexpected left-edge prosodic boundary tones to words where a right-edge boundary tone was expected, Roll & Horne (2011) found an early (N100) anterior negativity. Eckstein & Friederici (2006, p. 1703) also find a negativity for syntactically correct sentences with prosodic violations, however, this effect occurred later at around 300-500 ms. Since the target words where bisyllabic and the violation occurred in the second syllable of the word, this divergence from the other results might be related to the latency of the violation. It should be kept in mind that these studies created the expectation of a prosodic contour through syntactic cues. In the study that will now be described, it is the semantic prediction generated by sentence context that creates an expectation for the prosodic contour.

Chapter 3 The study

This chapter describes the EEG study that was conducted to find an answer to the research questions posed in the introduction. Before going into technical details, it incorporates the insights from the literature review to further specify and operationalise the independent variables and some other terms that the research question includes. Then, it describes the recruitment criteria for the participants and gives some information on their biographical data. Following this, the chapter gives an account of the material that was required for the study and how it was created and designed. It puts a special focus on the acoustic characteristics of the stimuli, since this is the basis of the effect that was investigated. After this, both the procedure in the lab itself is described and details for the EEG recordings are given. Lastly, the chapter describes the processing methods and statistical procedures that were applied to the collected data and describes ethical considerations.

3.1 Operationalisation

The primary independent variables that the research question mentions are the predictive power of the sentence context and the realization of the prosodic contour. Furthermore, the research question specifies the dependent variable by stating that it will measure the perception of the contours. The hypotheses then go further in specifically stating that the perceptive saliency should be modulated by the independent variables. These are the primary variables whose interaction the study investigates. Furthermore, two behavioural variables will be measured to control participant attention while listening, which are the response accuracy and reaction time to comprehension questions during the listening phase. All variables are elaborated upon in this subchapter.

While mismatch paradigms create expectations locally through the repetition of a stimulus, many studies that explore linguistic prediction modulate the prediction through a preceding sentence context. This is meant to guarantee ecological validity. To ensure that a word is predicted, cloze probability tasks are often utilized. Before conducting the EEG study, the stimulus sentences are tested with different participants to evaluate the cloze probability of the target words. Based on these results, sentences with target words that do not meet the cut-off probability are removed from the data.

Due to time constraints, this process was not feasible for the current study, which is why a different method was used to ensure high predictability. All target words were adjectives and placed at the end of a sentence. The sentence contexts with high predictive power, from now on called “highly constraining”, were created by two methods. One method was juxtaposing pairs of opposing adjectives, one being the target word (see example (1)). Another method was associating adjectives with words that were closely semantically related, as in example (2). The sentence contexts with low predictive power, from now on called “lowly constraining”, were created by ensuring that the sentence had many possible feasible completions, as in (3). More information on how exactly the sentences were constructed can be found in subchapter 3.3.1.

(1) Bilen var inte långsam utan den var snabb.

The car wasn’t slow but it was fast.

(2) Han var tvungen att värma upp maträtten för den var kall.

He had to warm up the dish because it was cold.

(3) Mattorna passade inte till inredningen för de var röda.

The carpets didn’t fit the furnishing because they were red.

The realization of the prosodic contour is also slightly more complicated than creating a simple word or phoneme mismatch. Previous studies on word accents have created mismatches by switching the word accent contours. However, since this study aimed to disentangle word-level processes from more basic prosodic processes, this was not suitable because it would engage word-level predictive processes as seen in those studies. Instead, the anomaly was created by choosing a contour that wasn't mismatching, but unusual. Due to how the sentences were constructed, all adjectives contained new information and therefore were supposed to be realized with the focal version of the respective word accent (Krifka, 2008). For Accent 1, the anomaly was the non-focal version of Accent 1 when the focal version was expected. For Accent 2, it was a delay of the first accentual peak followed by the regular focal peak, resulting in a contour that doesn't occur in Stockholm Swedish. More technical details on how this was done can be found in subchapter 3.3.2.

As many previous studies, this study also uses the EEG/ERP methodology. More specifically, it mainly focuses on early negative components in the grand-averaged ERP waveforms. These have been linked to mismatch responses in MMN and PMN studies, as well as studies on prosodic pitch processing, as explained in the literature review. However, since this study uses a novel approach, no specific electrode regions of interest are pre-defined for the statistical analysis. A higher saliency as formulated in the research question is operationalized as a higher amplitude of the negative component. As already introduced above, the sentences that the participants listen to are interspersed with comprehension questions to control for attentional effects. The accuracy and response time of the participants will therefore act as two further dependent control variables to supplement the main dependent variable of ERP component amplitude.

3.2 Participants

17 native speakers of Swedish who grew up in the Stockholm area were recruited for the experiment. Since it is very hard to recruit monolingual speakers in this age group, information on other languages spoken was collected. English was spoken by all participants, but eleven also spoke other languages, including German (n=4), Spanish (n=4), Persian (n=2), French (n=2) and one each of Norwegian, Greek, Latin, Russian, Finnish and Czech. All the participants were right-handed, had no neurological conditions and no hearing impairments or visual impairments

that hadn't been corrected with glasses. Twelve of the participants were men and five were women. The participant's age ranged from 20-28 years, mean age was 23,7 years (SD=2,88). The participants were recruited through word of mouth, flyers and internet advertisements through the portals Accindi and Reddit. The flyers and the text used in the advertisement can be viewed in the appendix.

3.3 Materials

3.3.1 Sentence frames

A total of 160 sentence frames were generated for the experiment based on 40 monosyllabic adjectives. Each word occurred in four of the sentence frames, twice in singular and twice in plural. Since all monosyllabic words have Accent 1 in Swedish and all plural adjectives are disyllabic and are assigned Accent 2, each word therefore also occurred twice in Accent 1 and twice in Accent 2. Each singular and plural adjective occurred once after a highly predictive and once in a lowly predictive sentence frame while still being semantically congruent. To ensure that the prosodic environment of the target words didn't deviate between conditions, the sentence frames were constructed so that only the first part of the sentence changed, while the latter stayed constant in all conditions (except for the plural inflections). Table 2 illustrates the four sentence frames with the word "tung" (heavy). High predictiveness of the sentences was ensured with pairs of adjectives that were opposites.

Table 2: The sentence frames

The four different sentence types in the stimulus material Including English translation			
Condition	Context sentence	Constant sentence	Translation
SG-HI	Jag antog att resväskan var lätt men	den var tung	I assumed the suitcase was light but it was heavy
SG-LO	Jag tog inte med mig boken för	den var tung	I didn't take the book with me because it was heavy
PL-HI	Vikterna var inte lätta,	de var tunga	The weights weren't light, they were heavy
PL-LO	Vad han inte sa om dörrarna var att	de var tunga	What he didn't say about the doors was that they were heavy

All adjectives in the four different sentence frames can be found in the appendix. When choosing the adjectives, care was taken that none of them included voiceless consonants in the medial or final positions so that the f0 contour was not interrupted.

The sentences were recorded in a sound-isolated booth with a microphone that was at a fixed distance to the speaker (15cm). They were spoken by a female native speaker of Swedish who speaks with a very distinct Stockholm accent. All recordings were captured into on channel with a high quality condenser microphone at 44,1 kHz in the software Sopran (Granqvist, 2009). The speaker was instructed to keep her loudness and speech rate as constant as possible. Since the sentences were highly regular, the speaker did not have to make a strong effort to keep the prosodic realizations regular, which also attests to the naturalness of the prosodic contour present in the recordings. All sentences were recorded three times, but two of those instances were unfortunately lost to a technical malfunction at the recording stage.

3.3.2 Creation of the anomalous contour

When deciding on how to manipulate the target words so that they included an anomalous contour, several considerations had to be made. The most straightforward way would have the speaker produce a version of the constant sentence with a varied prosodic contour on the target word. However, this could have led to several problems. First, since the contours are unnatural, it would have required the speaker to actively monitor her prosodic contour and keep the manipulation constant. This was judged very difficult considering the large number of sentences to record. Second, as already discussed in the literature review, Swedish word accent contours are not limited to the word they occur in but interact with the words around them. If the speaker had produced an anomalous variant of the target word, the words leading up to the target word could have contained prosodic cues to the realization of the target word. In prior experiments, these cues that do not correspond to the typical contour have resulted in very early negativities in the ERPs that might be confounded for an early mismatch effect (Nieuwland, 2019, p. 371; Zwitserlood, 2004).

Therefore, all prosodic manipulations were done afterwards in Praat (Boersma & Weenink, 2022). This ensures that the anomalous contour is truly the only difference between the conditions and that there are no cues to a different contour leading up to the target word. It needs to be said, however, that, in addition to changing the intonational contour of a word, the manipulation gives some words a slight artificial quality, especially on high rounded vowels. This needs to be taken into account when discussing the results.

The recordings were processed and annotated in Praat (Boersma & Weenink, 2022) according to the following procedure: All sentences were cut into two sections, the sentence before the constant sentence conclusion (called "context sentence" from here on) and the constant sentence conclusion itself (called "constant sentence" from here on). To ensure that only the context sentence deviated between the low and high constraint conditions, one of the two versions of the constant sentence that existed of every target word form was chosen to be played after both the highly and lowly constraining context sentences. The selection of the constant sentence was mainly based on the realization of the target word since some acoustic traits make intonational contour manipulation in Praat difficult. For example, care was taken that none of the target words in the chosen constant sentences contained creaky voice. Additionally, the time point for cutting the sentences in two sections and reattaching the chosen constant sentence to both the highly and lowly constraining sentence context was chosen so that the cut was as imperceptible as possible.

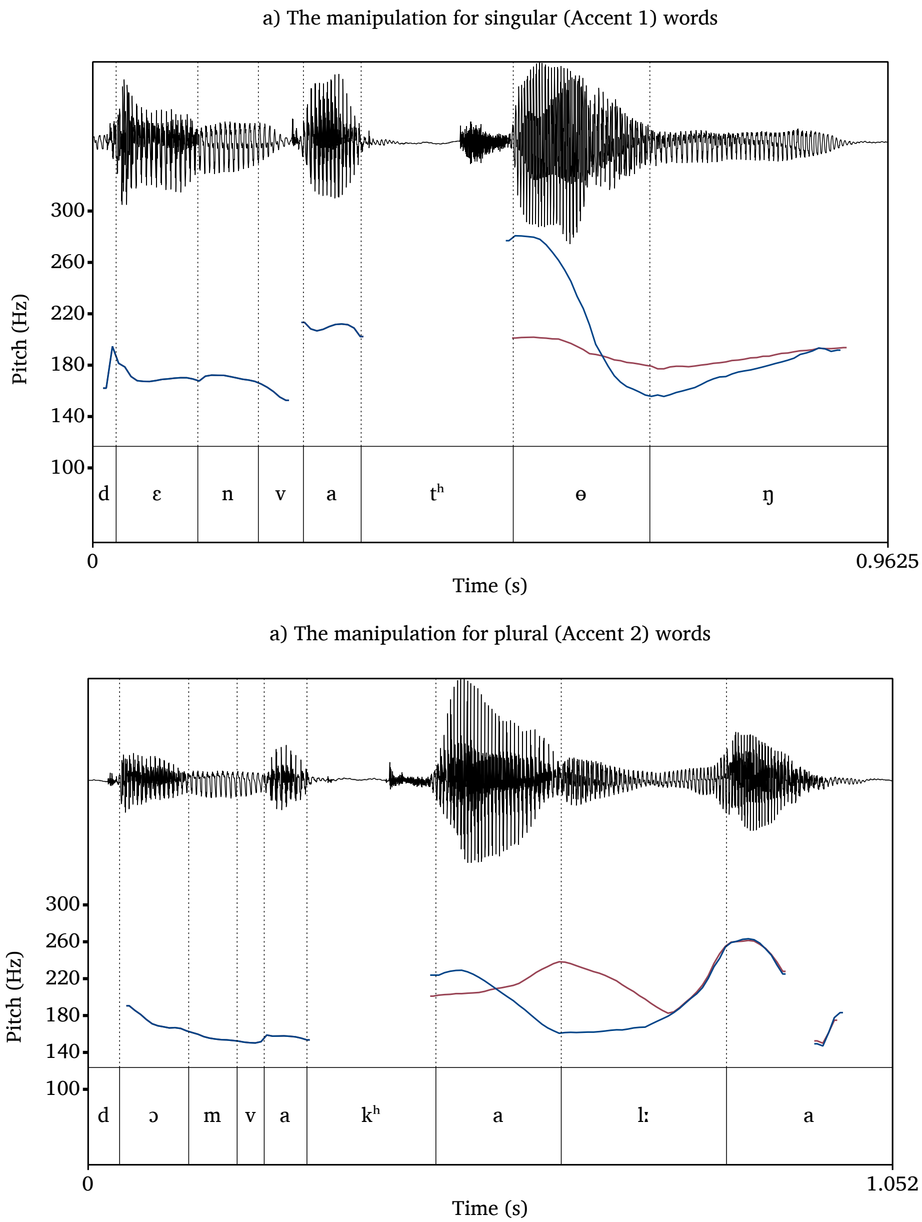


Figure 3: The manipulations of the F0 contour for singular and plural words

To create the anomalous contour, every target word's phonological contour was with Praat's overlap-add manipulation. For the Accent 1 words, the accentual peak was replaced by a low tone. For the Accent 2 words, the onset of the first peak, which is normally aligned with the onset of the stressed syllable in Stockholm Swedish, was pushed backwards to be aligned with the coda of the stressed syllable. An example of the normal and the manipulated F0 contour for singular and plural words can be seen in Figure 3. This created a percept that was judged unnatural by several Swedish speakers. To illustrate the differences in the f0 contour between the typical and anomalous conditions, bulk measurements of the f0 landmarks were taken and are displayed in Table 3.

As evident from the table, the F0 peak was significantly delayed for the plural words in the unexpected version, while the actual range of the pitch stayed similar. The words with anomalous contours contained a slightly smaller pitch range on average, which was due to the fact that for some words, maintaining the exact pitch height in the delayed condition led to very unnatural-sounding words. For the singular words, the table only reports the pitch range since there was no determinable peak in the anomalous conditions. However, it is clear from the results that the pitch range over the whole word was quite reduced. In fact, the most likely source for a bigger pitch range in the words was a slight upwards contour at the end of the syllable which occurred naturally in some words and was left untouched by the manipulation.

Table 3: Manipulation effects

Effect of the manipulation on F0 characteristics
Shown for singular (Accent 1) and Plural (Accent 2) words

		Normal		Anomalous	
Measurement	Measurement	Average	St. Deviation	Average	St. Deviation
Plural	Peak Timing	200.3878	44.68749	334.46825	46.46108
Plural	Pitch range	122.0414	19.24004	112.41105	16.69957
Singular	Pitch Range	132.4044	57.45296	44.37139	23.44745

3.4 Procedure in the lab

Once the participants had arrived at the lab, they received an explanation of how the experiment would proceed and what the participation entailed from their part. They were also informed that their personal information and the experimental data would be handled and stored in accordance with the General Data Protection Regulation (GDPR) of the EU and the Swedish Data Protection Act. Additionally, they were informed of their right to cease participation at any time and to request the deletion of their data to the extent that it is possible at that point. After this, the participant and the experimenter signed a consent form and the participants completed a questionnaire that collected personal data such as age, place of birth, native language(s) and primary residence during childhood. Both the consent form and the questionnaire can be found in the appendix. After the experiment was concluded, the participants were rewarded for their participation with a movie ticket corresponding to about 100kr.

The experiment itself consisted of the stimulus sentences interspersed with comprehension questions to control that the participants were attending to the stimulus. The participants were introduced to the procedure in two practice trials. Before each auditory presentation, there was a second of silence while the fixation cross was displayed. Then, the sentence played in its entirety. The participants were instructed to try not to blink during the auditory presentation and were asked to blink or adjust their posture during breaks. The breaks were scheduled through pause sequence that was initiated after every ten sentences. The pause screen encouraged the participant to take a small break. Therefore, the sentences were presented in sixteen blocks of ten sentences each.

In irregular intervals, a comprehension question was displayed on the screen after a sentence had been presented. All questions were related to the whole meaning of the sentence and not just the target word to avoid making the focus of the experiment apparent. The questions generally adhered to the sentence frame “Var det [noun] som var [adjective]” (Was it the [noun] that was/were [adjective]), a full list of the questions can be found in the appendix. The participants could respond to by pressing a green or red key on the keyboard. After answering the question, the next sentence was presented automatically. In the case where no question was given, the sentence presentation was self-paced, meaning that the participants pressed the space bar to continue to the next sentence. There were a total of 61 questions included in the experiment, 31 of which had the correct answer “yes” and 30 of which had the correct answer “no”. Additionally, the questions were balanced as to whether they occurred after a normal or an anomalous contour. Whether one question more occurred after the typical or the anomalous prosodic contours was counterbalanced across the participants.

3.5 EEG recordings and processing

To prepare the experiment, the participants were fitted with an EEG cap that corresponded to their head circumference. In addition to 32 electrodes that were placed according to the international 10-20 System (Homan et al., 1987), two electrodes were placed above and below one eye and two electrodes were placed on the temples to account for eye movements. For later re-referencing, an electrode was placed on the mastoid bone behind each ear of the participant. The ground electrode was attached to the forehead of the participant. Impedances were kept below 5 kΩ for the scalp electrodes, 6 kΩ for the eye electrodes and 2 kΩ for the average mastoids. The EEG signal was sampled at 100 Hz for the first twelve participants and at 250 Hz for the last five participants with the software Curry 7 Neuroimaging Suite (Compumedics Neuroscan).

The EEG data was pre-processed using the EEGLAB software (Delorme & Makeig, 2004). First, any prolonged stretches of EEG data where the participants were talking to the researcher or taking a break were deleted and an offline 30 Hz low-pass filter, as well as an offline 0.1 Hz high-pass filter were applied to the data. The EEG data was then visually inspected and any excessive muscular artefacts were manually removed from the data, except if they occurred close to or during an EEG timestamp. Afterwards, an independent component analysis (ica) using the runica algorithm was conducted to identify blinks, eye movements or other artifacts in the data that would contaminate the results. The three most common eye artifacts were then removed.

The EEG files that were sampled at 250 Hz were then downsampled to 100 Hz for further processing.

Finally, the parts of the EEG signal that were relevant to the analysis were identified using the EEG time-locking points that were sent to the EEG computer during the acquisition. Around the time-locking points, a window of 200 ms before and 800 ms after the points was extracted for analysis and the epochs were baseline-corrected using a baseline period of 200 ms before the time-locking point. Then, any epoch containing activity that exceeded 100 µV was excluded from the analysis since this generally does not stem from EEG activity. In total, this occurred during 15 epochs.

3.6 Statistical analysis

3.6.1 Behavioural results

The results files created by Psychopy were preprocessed in RStudio (RStudio Team, 2022) using the tidyverse, gt and ggplot packages. The descriptive results were compared in the four conditions created by the two independent variables context constraint and prosodic contour. For the reaction times, the effect was further investigated using linear mixed models, which were implemented with the lme4-package (Bates et al., 2015). Linear mixed models are preferable to generalised mixed models since the experiment collected reaction time data in a within-subjects design. This violated the independence assumption, meaning that there are data points in the data that are related by virtue of belonging to the same participant. Additionally, linear mixed models have the advantage over linear regressions that they allow for random effects with varying slopes and intercepts, which is often needed to describe experiments with fatigue or familiarization effects as well as differing performance baselines of the participants (Baayen, 2008, pp. 263–327; Winter, 2020, pp. 232–244).

3.6.2 Electrophysiological results

The stimuli sentences exist in four conditions as explained in subchapter 3.3. Additionally, each of the conditions exists in a typical and an anomalous version, bringing the total number of conditions that are analysed in this study to eight. For each of these eight conditions, a grand average waveform was created in Matlab. However, not all conditions are compared against each other: The results for Accent 1 and Accent 2 will be considered separately in the results and the discussion because of the syllabic difference between the words. For every word accent type, the differences of the amplitudes for the typical and the anomalous contour in the low and high constraint condition will be compared using a non-parametric approach.

In general, the statistical analysis of EEG data suffers from the multiple-comparisons-problem. Since many electrodes are sampled at many time-points, not correcting the significance threshold leads to an unacceptable family-wise error rate (FWER) and false positives (Groppe et al., 2011). At the same time, corrections like the Bonferroni correction tend to be very conservative when applied to EEG data, leading to p-values that require very large effects to reach significance. Therefore, experimental approaches that don't pre-define the region of interest can benefit from non-parametric statistical methods such as permutation analyses. In the case of this experiment, a cluster-based permutation analysis was chosen that can detect significant clusters of electrodes and time points in the data for a constrast.

To conduct the permutation analysis, the Fieldtrip Toolbox (Oostenveld et al., 2011) was used. Specifically, cluster-based permutation analyses using the Monte-Carlo method were conducted for the time of 100 to 800 ms after the time-locking point. If there were any clusters detected, their significance was determined using an alpha level α=0.05. The conditions were compared by substracting the results of highly constraining sentence contexts from the result for lowly constraining sentence contexts for each of the prosodic contours and then comparing them with each other.

3.7 Ethical considerations

Since the study at hand recruited human participants and obtained biographical as well as physiological data from them, some ethical considerations are in order. Of course, all usual data processing and anonymity standards were adhered to, and the participants were compensated for their help, as described in the previous subchapters. However, beyond that, care was taken that the participants also felt like their help and that their experience during the EEG was valued. Since an EEG is also not a very usual procedure, they were introduced to the methodology and the basic tenets of the technology, and all preparatory steps were explained to them. Once connected to the EEG system, they had the opportunity to look at the recording and see what effects movements or blinks had. After they had completed the experiment, they were also able to look at alpha waves in the signal if they occurred during the EEG session.

Furthermore, the participants had to be kept uninformed of the exact purpose of the experiment until the EEG recording was over. While it would have been preferrable if they had more detailed knowledge of the research questions, an awareness of how and when they phonologically anomalous contours would occur would have influenced the results of the study. Therefore, the participants were merely told that the study investigated “sentence comprehension”. After the EEG recording was over, the participants were asked about any regularities or anomalies in the sentences and subsequently informed about the purpose of the study and the study design. Any questions that they had about the study were also answered as far as it was possible.

4 Results

This chapter reports the behavioural and neurological results of the EEG study which was conducted to answer the research questions posed at the beginning of this essay. First, it focusses on the control measures and shows the reaction time and accuracy when answering the comprehension questions during the EEG session. Afterwards, the electrophysiological results are presented separately for singular (Accent 1) and plural (Accent 2) words.

4.1 Behavioural results

A limited number of sentences were followed by a comprehension questions. The participants were not instructed to reply as fast as possible, since the priority in the experimental procedure was to avoid blinking and other artefacts to contaminate the data. Therefore, beyond controlling for attention, the accuracy and reaction time data may have limited value for understanding how the participants reacted to the stimuli. Unfortunately, one participant's behavioural data was lost due to an error in closing the Psychopy experiment file. This reduces the number of participants for this dataset to 16. Furthermore, one subject switched the response keys so that they pressed the "no" answer nearly every time the correct answer was a "yes" and vice versa. This was accommodated when calculating the response accuracy.

4.1.1 Accuracy

Table 4 displays the mean response accuracies divided by the context constraints and the prosodic contour that they followed. There was a relatively high accuracy overall, probably owing to the unrushed nature of the experiment design. However, it seems that questions following highly constraining sentence contexts were more often answered incorrectly. The lowest accuracy occurred after highly constraining sentences and normal prosodic contours, but the accuracy for highly constraining sentences followed by anomalous contours was also lower than after lowly constraining sentences. The data does not show a trend as to whether normal or anomalous contours affected the accuracies in a particular way. In highly constraining sentence contexts, the questions following normal contours were less often answered correctly, while they were more often answered correctly in the lowly constraining sentence contexts.

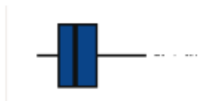
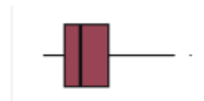
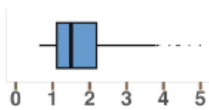
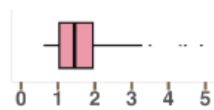
Table 4: Descriptive results of Answer accuracy

Comprehension question answer accuracy Separated by context constraints and prosodic contour		
	Prosodic contour (in %)	
Context constraints	Normal	Anomalous
High	87.12	91.25
Low	96.90	96.40

4.1.2 Reaction times

As mentioned, the reaction time data must be evaluated with caution, because the participants were not instructed to answer as fast as possible. This can be clearly seen by the many outliers in the data and the generally slow reaction times. However, this is expected considering that the participants read and understand an entire sentence before replying. Table 5 contains the mean values of the reaction times divided into those following typical and anomalous prosodic contours. Additionally, the table contains boxplots that give some information about the skewness and spread of the data. Note that for better visibility, all outliers above 5 seconds were removed from the boxplots.

Table 5: Descriptive results of reaction times

Comprehension question answer reaction time Separated by context constraints and prosodic contour						
	Typical contour			Anomalous contour		
Context constraints	Mean (s)	Standard Dev. (s)	Quartiles	Mean (s)	Standard Dev. (s)	Quartiles
High	1.940358	1.266634		2.006215	1.720110	
Low	1.811539	1.109216		1.740731	1.116254	

The results show that the mean reaction time did not change following anomalous contours, although the data for anomalous contours is slightly more skewed to the right for the highly constraining sentence contexts. Instead, the context constraint of the sentence frames seems to be a factor in the reaction times. The medians and means for lowly constraining sentence contexts are slightly lower than those of the highly constraining contexts. Additionally, the boxplots show more spread for the highly than the lowly constraining sentence contexts.

When trying to understand the reaction times, fatigue and familiarization effects need to be taken into account. To explore whether the trial number should be considered when modelling the effect of the independent variables context constraint and prosodic realisation on the reaction times, Figure 4 shows the response latency as a function of the trial number of the sentence. Response times over seven seconds are not shown in the figure to make the scatterplots more easily readable. The participants seem to differ as to how their response times changed over the experiment, with some participants showing familiarization effects, while others show fatigue effects. This is a good indicator that the trial number should be included as a main effect in the model.

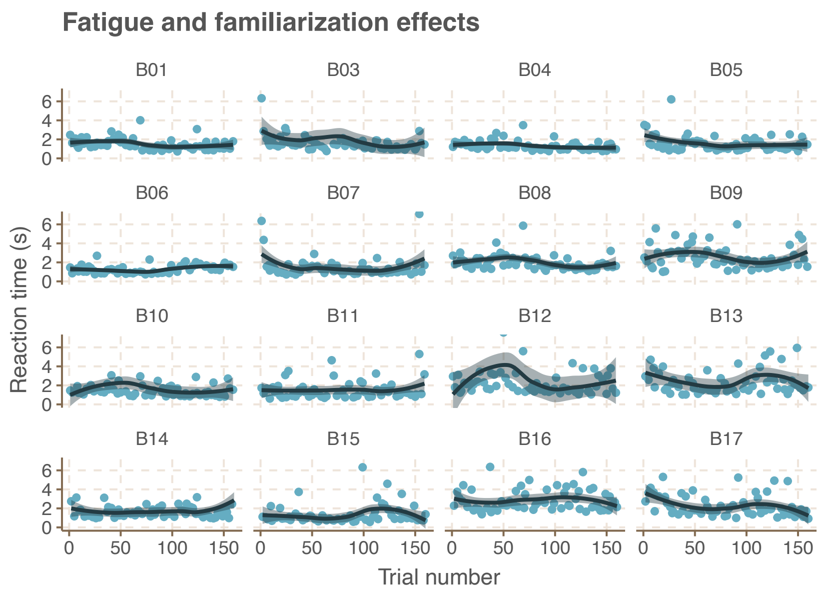


Figure 4: Fatigue and familiarization in the reaction time data

4.1.3 Statistical analysis

The descriptive statistics showed some differences for the variable "sentence constraint". However, since response times and accuracies are highly variable between participants, linear mixed models were used to explore if the differences between the conditions were statistically significant. This allows us to get a closer understanding of the data while factoring in individual participant differences as well as the familiarization and fatigue effects present in the data. The independent variables of sentence context, prosodic contour and trial number were included in the models as fixed effects, together with the random effect of participant. To see which combination of fixed effects provided the best model fit, four different models were tested. Table 6 provides an overview over the models and significance levels for the fixed effects. Additionally, it shows the overall goodness of fit split up for only the fixed effects and for both the fixed and random effects as seen through marginal and conditional R^2-values respectively. Since the fixed effect of trial and the random effects were always included in the model, the model description only shows the configuration of the fixed effects corresponding to the independent variables of the experiment.

Table 6: Results of the inferential statistics

Model outcomes and fits Results of linear mixed models fitted to reaction time data					
	Model fit		Model outcomes		
Model name	R^2m	R^2c	df	χ2	p
Full					
Sentence context * prosodic realisation	0.0117	0.1354	965.3	0.1623	0,6713
No interaction					
Sentence context	0.0115	0.1354	960.0	-2.3850	0.0173
Prosodic realisation	0.0115	0.1354	960.1	0.1050	0,916
Only context					
Sentence context	0.0000	0.1303	960.0	-2.3980	0,0167
Only prosody					

The R^2-values show that the effects included in the models account for around 13% of the variance, however, this is mainly due to the random effect of participant. The fixed effects have a significantly lower contribution to the model fit, especially the prosodic realisation. However, the sentence context does seem to be a better predictor, as also visible in the boxplots. The p-values show that the reaction times after highly predictable sentences are significantly higher than after lowly constraining sentence contexts. This means that it took the participants longer to respond to a comprehension question after a highly constraining sentence, irrespective of the prosodic realisation of the target word, confirming the impression gained by the descriptive statistics.

4.2 Results from the EEG recordings

Since this thesis tests the interaction of two independent variables, four grand average waveforms of the ERP data must be compared. Additionally, since half of the target words contained Accent 1 and half contained Accent 2, the grand average waveforms for these should be considered separately. Therefore, this subchapter covering the results of the ERP study proceeds as following: For each word accent the grand average waveforms of all four conditions is shown, as well as the grand averages for the highly constraining and lowly constraining sentence contexts separately. This allows a clear overview of the effect the anomalous contours had on the ERPs in both conditions. The statistical subchapter proceeds in the same manner as the subchapter showing descriptive waveforms. Not all electrodes that were used to record the EEG data are shown in the grand average waveform compilations, however, all are included in the statistical analysis.

4.2.1 Accent 1

In the conditions with an anomalous contour, the target words were realised with a low tone rather than a high tone. However, a high tone, the focal version of Accent 1, would be expected since the target word introduces new information. Figure 5 shows the grand average waveforms of the event-related potentials in the four experimental conditions. The anomalous contours are shown in red, while the typical realisations are shaded in blue. Additionally, the waveforms for words in lowly constraining sentence contexts are paler and the line is dashed. Negativity is plotted upwards in all graphs.

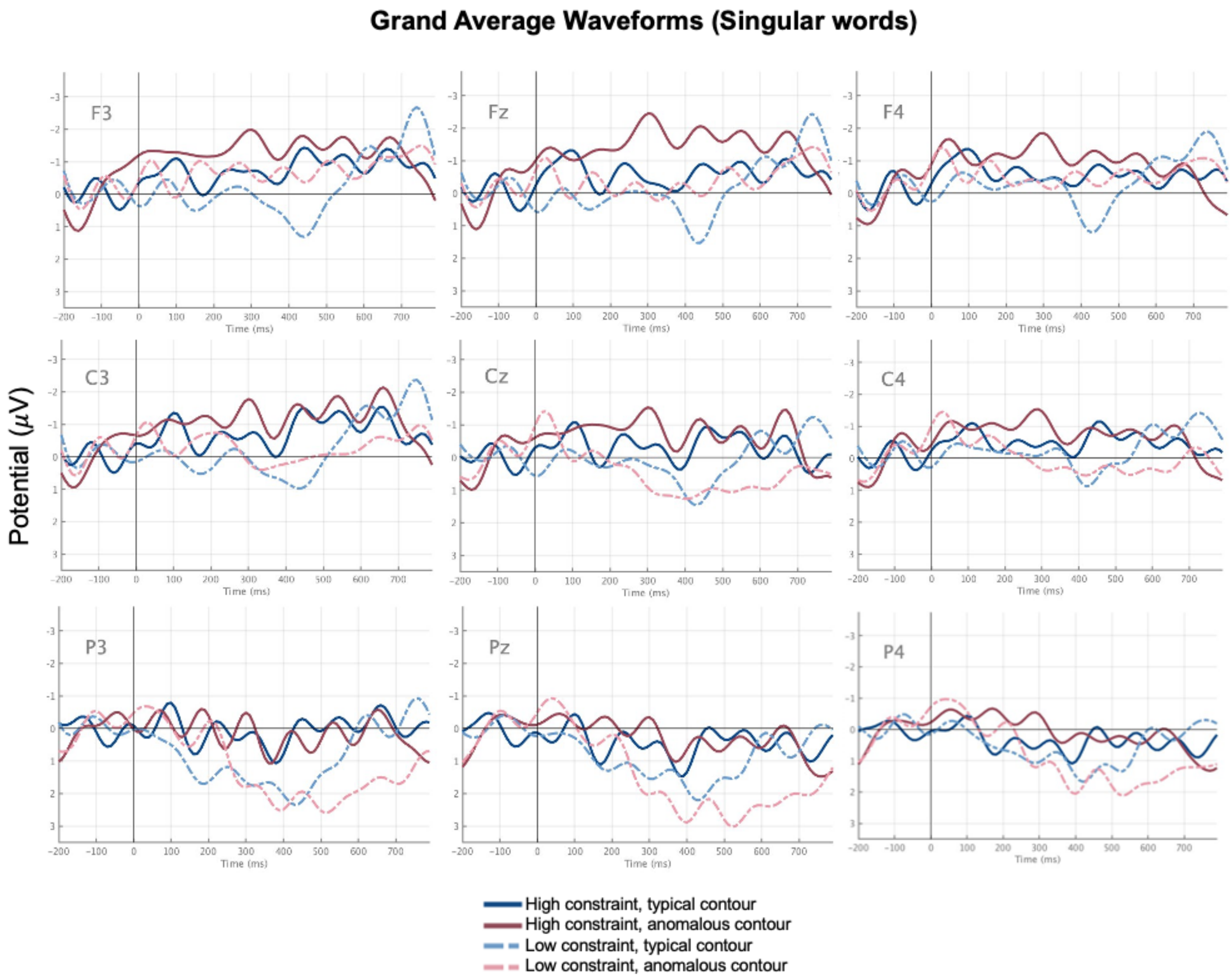


Figure 5: Grand average waveforms for singular adjectives, selected electrodes

There is an early frontal negativity that is present for the anomalous contours in the highly constraining contexts. This negativity seems to diverge at around 100 ms after the time-locking point, which is where the target words start to differ from one another. For most electrodes, especially the central and parietal ones, it rejoins the waveform of the typical target word realisations around 350 ms. However, it persists slightly longer in the frontal electrodes. All other grand average waveforms take a similar time-course up until about 350 ms after the time-locking point. Then, there is a quite pronounced positivity for the lowly constraining sentence contexts, especially in the posterior electrodes. This positivity lasts until about 700 ms, which is where the grand average waveforms converge again. After this, there is a positive deflection in the frontal and central electrodes when the pre-defined ERP-timescale ends.

To examine the grand average waveforms on the two levels of predictability separately, only the Cz electrode is shown in the next figures. There is less variance between the two realisations in the lowly constraining sentence contexts in the 100-300ms time range. However, in the late time range, there is a slightly bigger positivity in the condition with the anomalous contour. On closer inspection of the average waveforms of most electrodes, this difference is not consistent across all electrodes.

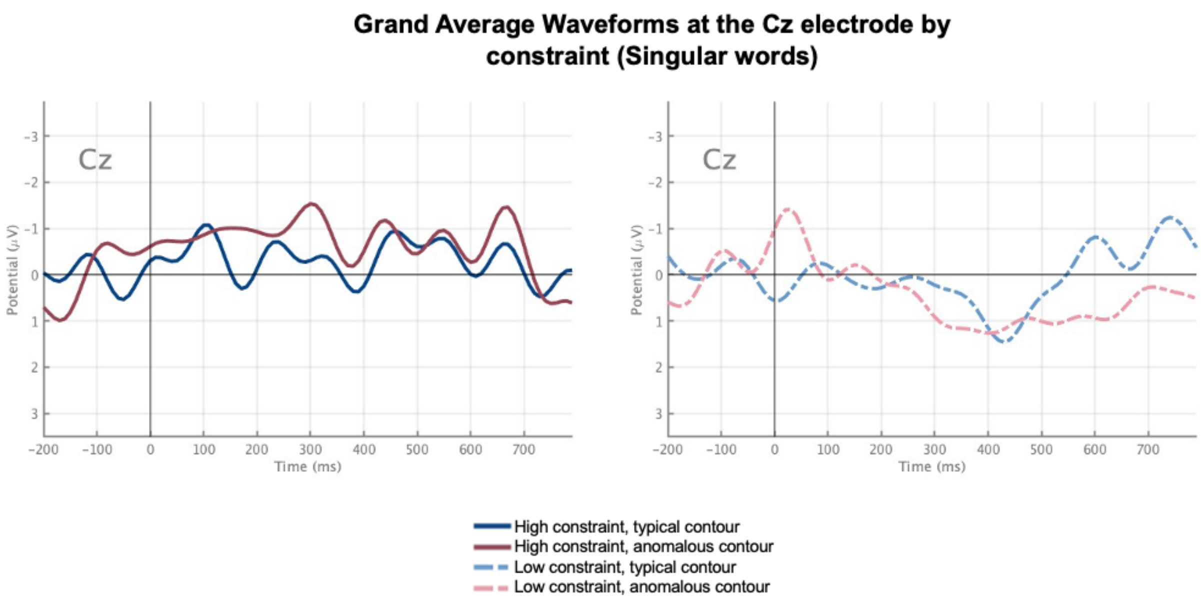


Figure 6: The differences in highly and lowly constraining sentence contexts at the Cz electrode for singular adjectives

Examining the waveforms following the typical and anomalous prosodic contours separately, the highly constraining sentence contexts lead to a more negative-going waveform overall when there is an anomalous prosodic contour. In these cases, the difference between a highly and a lowly constraining sentence context only become apparent after about 300ms after the time-locking point and it only persists for about 300 ms. The waveforms for the typical waveforms only diverge slightly at around 400 ms, when the low constraint condition is more positive.

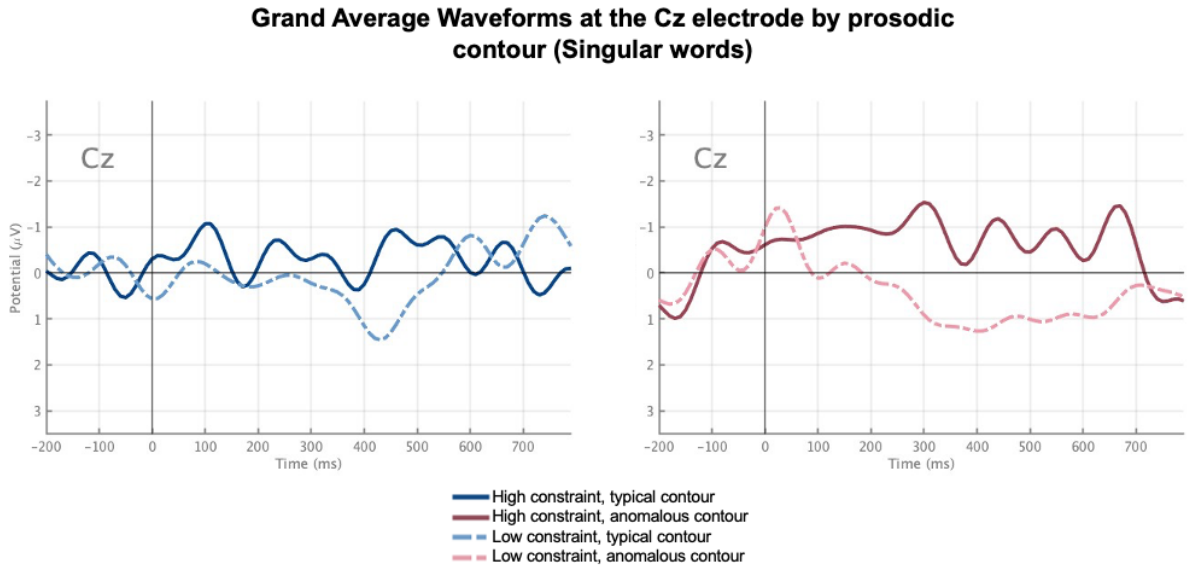


Figure 7: The difference between typical and anomalous prosodic contour, at the Cz electrode for singular adjectives

4.2.2 Accent 2

The manipulation of the f0 contour for the Accent 2 words led to a delayed onset of the first f0 peak. However, this peak was not reduced in amplitude. Figure 8 shows the grand average waveforms for all four conditions in these words. The larger negativity for the anomalous contour in the highly constraining sentence context is not visible in the same way here. Instead, the contour seems to be similar to the others up until the 300 ms mark, and there is a negative deflection for the typical contour instead. After this early period, waveforms for the typical and anomalous word accent contour in highly constraining sentence contexts seem to diverge. There is a larger positivity for the anomalous contours in the 400-600 ms range. The waveforms for the lowly constraining sentence contexts show less of a divergence, but show a similar voltage in posterior electrodes to the contour for the anomalous prosodic contour in the highly constraining sentence contexts. At the end of the pre-defined ERP-time-scale, there is a positive shift in the frontal and central electrodes, just like the one that was observed for the singular words.

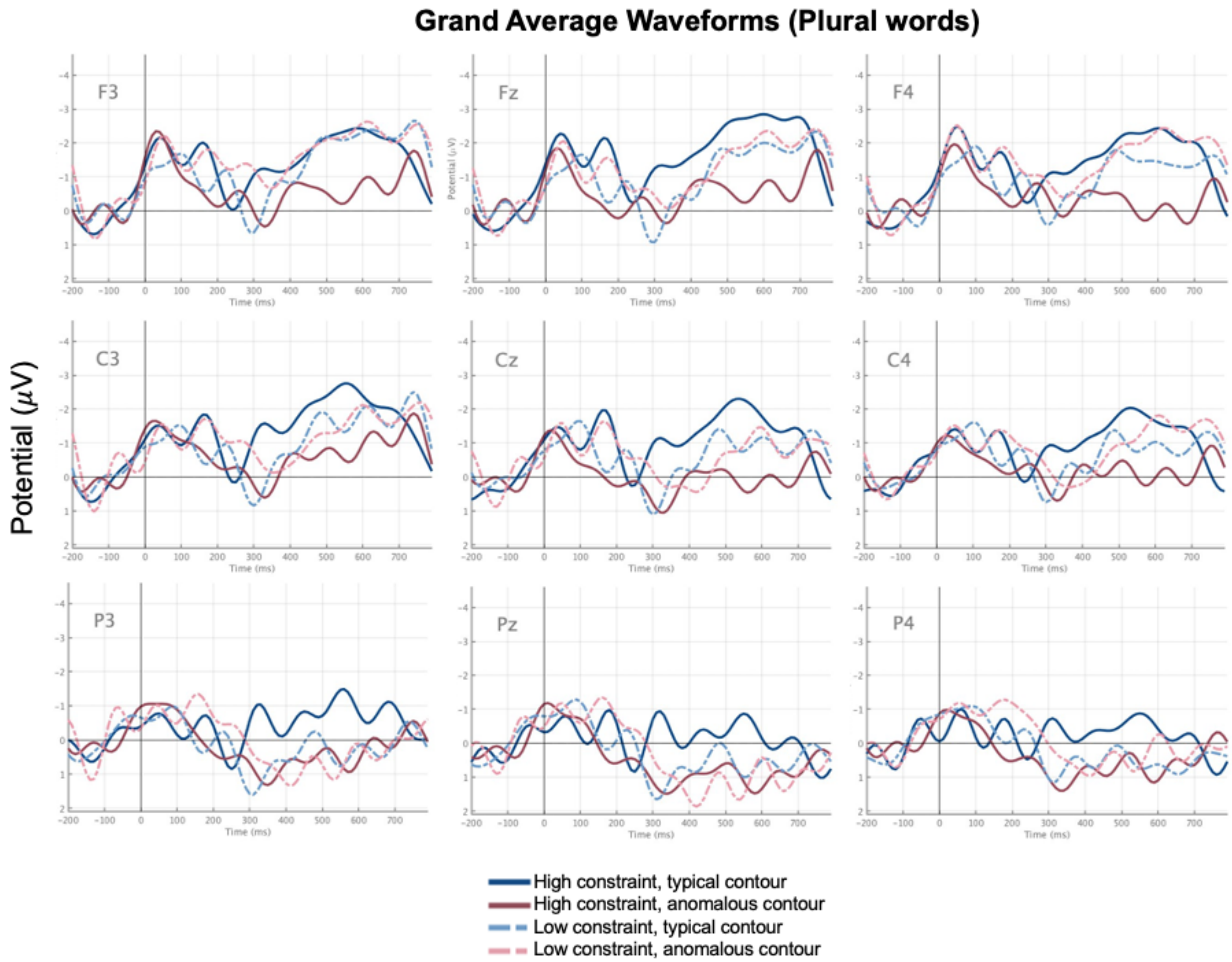


Figure 8: Grand average waveforms for selected electrodes for plural adjectives

Examining the ERP waveforms at the central electrode (Cz) separately for the sentence context condition, the larger positivity in the later time-frame is still clearly visible. However, it seems that there is an additional difference between the typical and anomalous contours in the earlier time frame. Contrary to the Accent 1 words, it is the typical contour that is eliciting a larger negativity. Examining the wave-forms, it seems that this negative-going deflection causing the difference occurs later for the anomalous contour. In the later time-frame, there is a clear difference in amplitude, with the anomalous contour causing a much more positive voltage. For the lowly constraining sentence contexts, the waveforms seem more similar, with only the negativity at around 400-500 lasting slightly longer before the waveforms converge again.

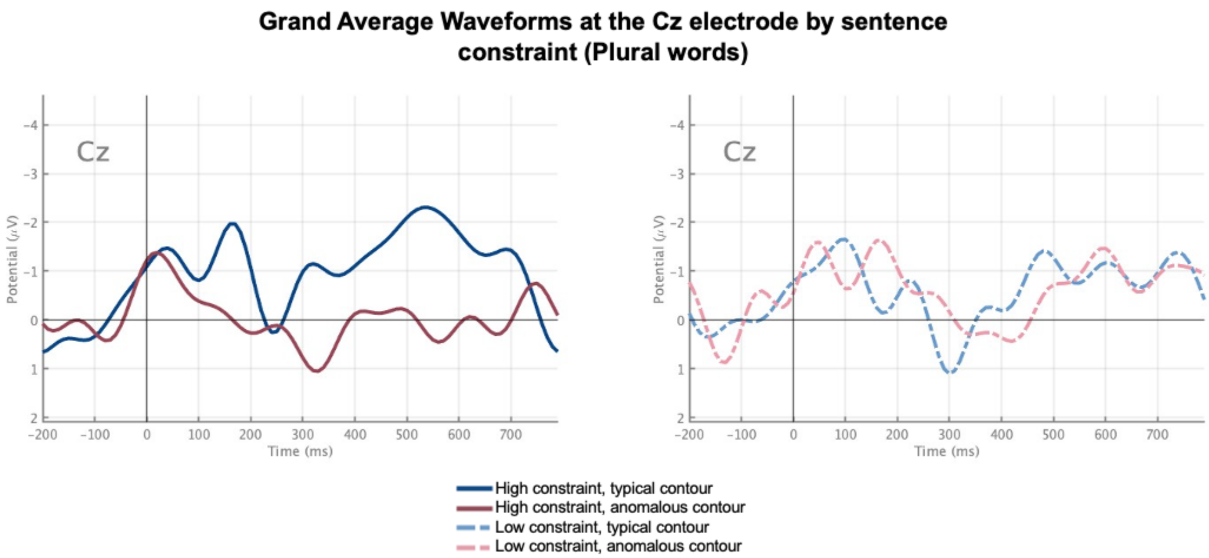


Figure 9: The differences in highly and lowly constraining sentence contexts at the Cz electrode for plural adjectives

The comparison of the waveforms for the anomalous and typical contours show a pattern that could be interpreted as a delay, which is visible for both typical and anomalous prosodic contours. This can be seen most clearly in the timing of the positive peak at around 300-400 ms. Secondly, is a difference between the waveforms in the later timeframe. While the highly constraining sentence frame induces a bigger negativity than the lowly constraining sentence frame in the typical contours, it shows a bigger positivity in the anomalous contours. Inspecting several electrode locations, this does not seem to be a generalized effect.

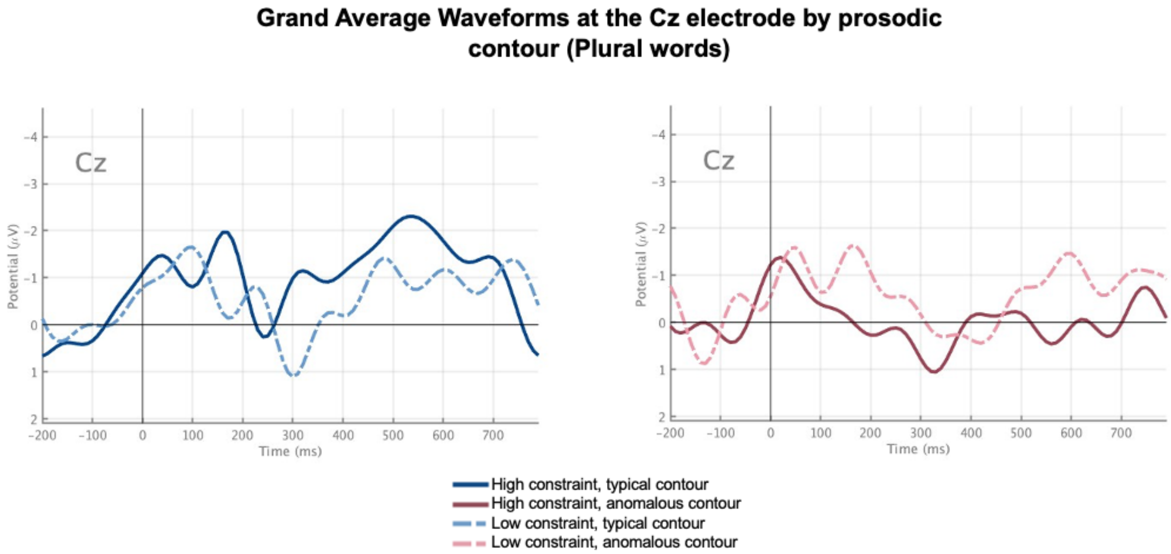


Figure 10: The differences between highly and lowly constraining sentence context at the Cz electrode for plural adjectives

4.2.3 Results from the permutation analysis

For the Accent 1 words, the permutation analysis found five positive clusters in the data. However, none of the clusters reached significance at an alpha-level of 0.05. The lowest p-value of the clusters was 0.2867 with a combined t-value of 41.0694. For Accent 2 words, the permutation analysis also found five clusters, however, they were negative. None of the clusters reached significance for this condition also, with the lowest p-value at 0.1538 with a combined t-value of -86.2686. Even when the time ranges where the clusters with the lowest p-values occured were examined through a post-hoc permutation analysis, there were no significant clusters that were found.

5 Discussion

This study aimed to provide insights into the way semantic contexts aided phonological predictions, and in turn how these predictions influence perceptual processes in listeners. Through an EEG/ERP design with the two independent variables sentence context and realisation of the prosodic contour, the study tried to disentangle the way the former modulates the perception of the latter. It was hypothesized that the sentence context would differently influence the perception in highly and lowly constraining context. The results of the behavioural data showed slightly less accuracies for highly constraining sentence contexts and a slower reaction time for comprehension questions following highly constraining sentence contexts. The ERPs results show an early negativity for the anomalous word accent contours in highly constraining, but not lowly constraining sentence contexts, but only for the words carrying Accent 1. The results for Accent 2 are less conclusive but show a late positivity for the anomalous prosodic contours and the target words in lowly constraining sentence contexts. This chapter discusses these results in regards to how they fit to the hypotheses, as well as the literature that was reviewed. Additionally, the chapter deals with limitations of the research and directions for future work.

5.1 Behavioural results

Concerning the behavioural data, it’s clear that the realisation of an anomalous contour did not affect the reaction times or accuracy in a significant way. However, there was a difference in reaction time and accuracy that was modulated by the sentence constraint. This knowledge is relevant for this discussion in two ways. First, it shows that the sentence as such did influence the reaction times. The results indicate that the sentences that were highly constraining led to slightly longer reaction times. While it is not possible to find a definite reason for this, it could be due to the contrastive nature of the sentences in the highly constraining sentence contexts. Since the sentences often contained two adjectives, it might have taken the participants slightly longer to remember which one of them was associated with the subject of the question.

Following from this, we can infer that comprehension of the sentences was not altered because of the prosodic contours. The participants seemed to have no problems answering the comprehension questions overall, and no significant difference could be found for the reaction times after anomalous prosodic contours. Any effect that these contours had are therefore most likely confined to word-level comprehension and did not affect the way the sentence in its entirety was understood. Therefore, it is less likely that the effects seen in the EEG are due to a difficulty to integrate the word into a larger sentence context.

What also needs to be considered is that the behavioural results, especially the reaction times, show a high variability, as seen by large standard deviations and relatively poor model fits. It can be assumed that a stimulus such as a sentence, especially a long sentence such as the ones used in this paradigm, influence the behaviour of the participants in many ways that can't be controlled. The sentences themselves were also quite varied, even if the sentence endings were kept constant. In addition to a prediction about the content of the next phrase, the participants likely also retained other details about the context sentence frame that influenced the perception of the following words.

5.2 Electrophysiological results

This perceptual process will be analysed more in depth in this subchapter, which describes the results of the EEG recordings. First, there are some general observations about the nature of the data that are important. The EEG data for the target words needs to be analysed with the fact in mind that the time-locking point was set to where the stimuli diverged, not to the word onset. This was the start of the rise in the first syllable. It is therefore possible that effects that would be usually later if word onset was used as a time-locking point appear a little earlier in the ERPs. Additionally, when discussing the results of the EEG recordings, it must be kept in mind that the conclusions are built on averages. While there might have been effects of the manipulations that are visible on average, this does not mean that lowly or highly constraining sentence contexts always lead to the observed effects.

Considering this, there are considerable differences between the grand average waveforms of Accent 1 and Accent 2. The results for Accent 1 show an early anterior negativity for anomalous prosodic contours in highly constraining contexts that is consistent with the early negativities observed for anomalous word forms in other studies (Bendixen et al., 2014), as well as the early negativities for anomalous prosodic contours seen in studies by Roll & Horne (2011), Eckstein & Friederici (2006) and Schön et al. (2004). Upon inspection, it seems that there is a bigger difference between typical an anomalous contour in the highly constraining conditions, although this could not be statistically validated with the permutation analysis. Therefore, the results need to be regarded as tentative. The fact that the negativity seems to be localized in more frontal regions of the brain also adheres to the results for anomalous words from PMN studies (Connolly & Phillips, 1994) and studies that apply a MMN paradigm to words (Boulenger et al., 2011).

The fact that it is the highly constraining context in which the additional effect is observed does not correspond to the results obtained by Brunellière & Soto-Faraco (2015), indicating that the results in that study might have stemmed from difficulties in word recognition. Instead, the results can be interpreted as showing an effect of prediction. The high tone that was not present in the anomalous prosodic contour is related to sentence focus. Additionally, it is tied to the word accent type of the predicted word. Therefore, it can be theorized that the participants expected the focussed version of Accent 1, building on a specific prediction about the upcoming word combined with knowledge about the sentence construction. The fact that the negativity was less pronounced for the lowly constraining sentence contour, even though the way the sentence was phrased made it very possible that a focally accented word was incoming, could mean that in absence of clarity on the upcoming word, no word form prediction is made and the listener relies on bottom-up cues.

It is interesting to note that this interpretation implies that there is no prediction of a focal accent, regardless of the word that carries it, when the semantic sentence constraints are low. At first glance, this would go against the results obtained by Eckstein and Friederici (2006), who didn't control for the predictability of the sentence but still obtained effects when a word-level prosodic contour was anomalous. However, they presented each sentence multiple times, making a third of the target words with anomalous contours highly predictable. It might therefore be the case that their results are due to the expectedness of the target word with its accompanying prosodic contour.

The late posterior positivity for the target words with Accent 1 in lowly constraining contexts are consistent with a P600 effect, possibly indicating increased processing difficulties due to a lowly constraining sentence context with many possible conclusions. The fact that this late positivity does not seem to be modulated by the prosodic contour of the target word strengthens the theory that it is related to sentence-level processes. It is also interesting that this late positivity goes against the results from the behavioural control measures, where comprehension questions following highly constraining sentence frames had longer reaction times. It might be that the constant sentence frame itself ('den var …') is easily processed because it is predictable, but the whole sentence is harder to understand because of the opposites.

The early negativity effect for anomalous contours that was described above was not obtained for Accent 2 words, where a very early negativity for typical contours in highly constraining sentence contexts was observed instead. Afterwards, the waveforms for the typical and anomalous contours in the highly constraining sentence contexts show a large deviation, with the anomalous contour eliciting a late positivity. In general, the manipulation that was chosen for Accent 2 words might have been less suitable for exploring the research question of this study, since the waveforms show similarities to results obtained in studies were Accent 1 and Accent 2 where switched (Söderström et al., 2017). Possibly, the manipulation of the Accent 2 words led the listeners to expect an Accent 1 word, which has also been shown to cause a P600. Since the late positivity is present in all conditions except the highly constraining sentences with a typical prosodic contour, the effect of the manipulation and the lowly constraining sentence context might be compounded so that no clear result can be reached.

Additionally, it could be the case that the listeners have some tolerance for the timing of the first peak of Accent 2. While the speaker of the sentences used in this study did not show much variability, the recording scenario was quite controlled, and more variety might be present in spontaneous spoken Stockholm Swedish than seen in the analysis of the F0 traits of the target words.

5.3 Overall discussion

Cycling back to the hypotheses, the inspection of the waveforms for Accent 1 show the most support for the Hypothesis H2, which states that anomalous prosodic contours are more perceptually salient in highly constraining than in lowly constraining sentence contexts. The hypothesis H1, which stated that anomalous prosodic contours are more salient in lowly constraining sentence contexts, receives less support from the EEG data, probably owing to the fact that the anomaly was too large to lead to phonemic restoration effects like the ones observed by Riecke et al. (2009).

However, these interactions of sentence context and prosodic contour do not reach statistical significance when conducting a Permutation analysis, as mentioned above. The hypotheses that the perception of anomalous prosodic contours is modulated by sentence constraints could therefore not be statistically validated and the data ultimately supports the null hypothesis. This does not mean that the H2 hypothesis needs to be rejected, as the inspection of the waveforms does confirm a trend that is in line with previous studies. Indeed, the presence of an early negativity as a response to phonological mismatches that does not reach significance mirrors the concerns voiced by Nicenboim et al. (2020) about the number of replication studies that observe trends, but no significant results. However, it means that the research question posed in the beginning does not receive a clear-cut answer through this study.

The differences between Accent 1 and Accent 2 additionally show that the type of anomaly in the prosodic contour has a direct effect on the results that are obtained. The interaction with word-level processes that was theorized in the previous subchapter seems especially critical. As soon as a bottom-up cue gives new information about possible word completions, it is utilized, as also demonstrated by van Petten et al. (1999). If therefore the manipulation also gives new information about possible word completions, the effect of the anomalous prosodic contour is confounded with processes pertaining to this new information.

Additionally, the difference between the results of Accent 1 and Accent 2 have implications for the level of detail present in the prosodic prediction. The results in this study could be interpreted in a way that the prosodic prediction mainly contains information about the absence or presence of a peak, not the extent of the peak or the exact timing. Considering the variability in spoken speech, this would seem economical. In fact, relating this back to predictive coding, it is coherent with the idea of precision weighting. If, for example, a rising contour is anticipated, the error signal might not be sent immediately upon encountering a level contour, but only when the level persists for longer than could be reasonably explained away by higher-level knowledge about the variability of the incoming input.

5.4 Limitations of the study

When considering the results, it is also important to keep in mind that there are some limitations to the current study design itself. Firstly, while the sample size is higher than in most early studies on phonological mismatches, it is still much lower than modern-day replication studies with more than a hundred participants. Additionally, whether the anomalous prosodic contours were clearly anomalous to Swedish listeners was not empirically validated in advance. Therefore, it can only be inferred from the results whether this was a contour that occurs within the natural variation of Swedish or if it was clearly not a typical contour. However, the fact that the F0-characteristics observed in the typical contours had a lower standard deviation than the difference between the typical and anomalous contours indicates that the anomalous contours are beyond the usual variation within typical F0-contours.

Another limitation is the creation of highly and lowly constraining context sentences since they were not validated with a cloze probability test. This means that the difference in expectedness is not statistically validated, but instead inferred from the way the sentences were constructed. Especially the pairs of opposites are expected to have led to high predictiveness, which was also commented on by many participants when they were asked about any observations that they made while listening. Additionally, it can’t be eliminated that some of the effect observed in this study was the result of priming due to the presence of the opposite adjective, although this is unlikely for several reasons. First, Otten and van Berkum (2008) show that the early negativities observed in their experiment which are similar to the ones seen in this experiment are not the result of priming, but dependent on discourse. Second, even if the word was activated by priming, this would not explain specific prosodic word form predictions that can only be made when the target word is considered in its larger, discourse-related context.

Lastly, the target words themselves were kept as uniform as possible, but this was limited by the fact that there are not that many monosyllabic adjectives in Swedish that only have one clear opposite. It is therefore likely that they were still varied enough to lead to a slight blurring of the effects in the EEG signal. For example, the timing of the F0 curve onset varied between target words since the sonority of the syllable onset was not controlled, When the target word started with a voiceless plosive like *kall* or a voiceless fricative like *fin*, the anomalous prosodic contour was only apparent at the vowel onset. For words with voiced onsets like *låg*, *ung* or *ny*, the anomalous contour started immediately. To control for the onset latency variation, the EEG time-locking points were placed at the points of F0 divergence between the typical and anomalous contours. This meant however that the time-locking points were not always set at the word beginning and the effects seen in the EEG signal might have been more pronounced if the stimuli were more uniform.

This is also related to a difficulty in understanding the effect of prosody in EEG settings as opposed to, for example, phonemes, which is that prosodic contours occur over a much longer time span. It is not clear from the data when exactly the participants noticed that the prosodic contour was anomalous in the context. This question has also not been well explored for Swedish, although there is some information about the phonetic implementation of Accent 1 and Accent 2 (Engstrand, 1995, 1997; Svensson Lundmark et al., 2017, see also subchapter 2.2.1) and there is some information on what is necessary to perceive a Southern Swedish Accent 1 or Accent 2 in the first place (Ambrazaitis & Bruce, 2006).

5.5 Further studies

The discussion has already raised several issues that could be explored in further studies. The differing results for Accent 1 and Accent 2 (or, to put it more generally, a pitch height and a timing violation), combined with the lack of clarity on the permissible variation for Swedish word accent contours, make it difficult to understand exactly how fine-grained prosodic predictions are. Further studies might supplement the results in this study by testing several different manipulations of the F0 contour for one word accent. This might come at the expense of only focussing on one of the two Swedish word accents to reduce the length of the EEG session. Additionally, this study should guarantee greater uniformity of the target words, possibly by making words predictable through other means than the sentence context. A computerized token test like the one used in D'arcy et al. (2000) might be suitable in this scenario.

Additionally, more clarity on the source of the effect could come from a higher electrode density or a magneto-encephalographic study. The source of the effects was generally not considered in this thesis, mainly because source-localising EEG data comes with some problems and would have required data from more than 32 electrodes. However, a source-localisation such as in Schön et al. (2004) could provide further confirmation that the negativity stems from the manipulated contour.

6 Conclusion

This thesis set out to test a core tenet of predictive coding, which is that predictable sensory input is processed differently than non-predictable input. In the case of the study, the sensory material in question was language, more specifically, prosodic contours. It was hypothesized that the predictability of words with anomalous prosodic contours modulated how they were perceived. To modulate this predictability, highly and lowly constraining sentence frames were constructed. The reaction to the target words in these conditions was then measured using the EEG approach while attention was controlled through behavioural measures.

The results show no significant differences between the perception of an anomalous prosodic contour in highly constraining and lowly constraining sentence frames, although there are some trends in the data that could be tentatively interpreted. Especially for the anomalous contour in the highly constraining sentence frame, a bigger early negativity at frontal electrodes could be identified in the waveforms. However, this was only true for manipulations that changed the pitch height, not the onset of an F0 peak.

The thesis therefore delivers some preliminary indications that sentence context, and in extension, the predictability of a word does modulate the perception of anomalous prosodic contours on the word. Additionally, it is the first study to consider the effect of anomalous prosodic contours while varying the predictability (not the predictedness) of the target word. The preliminary results also indicate that a focus on prosodic information could be a useful method when trying to understand the PMN and related phenomena. The discussion could also capture some of the phenomena that make prosodic word form prediction hard to study, such as the impreciseness of the sensory stimulus, difficulties with time-locking of longer-lasting phonetic units and the problems with confounding effects of phonetic prediction and perception with word-level effects of integration.

References

Abelin, Å., & Thorén, B. (2015). The relative perceptual weight of two Swedish prosodic contrasts. *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2015*. http://urn.kb.se/resolve?urn=urn:nbn:se:du-20794

Althaus, N., Wetterlin, A., & Lahiri, A. (2021). Features of low functional load in mono- and bilinguals' lexical access: Evidence from Swedish tonal accent. *Phonetica*, *78*(3), 175–199. https://doi.org/10.1515/phon-2021-2002

Ambrazaitis, G. (2009). *Nuclear intonation in Swedish: Evidence from experimental-phonetic studies and a comparison with German* [Doctoral Dissertation]. Lund University, Dep. of linguistics and phonetics, Centre for languages and literature.

Ambrazaitis, G., & Bruce, G. (2006). Perception of South Swedish Word Accents. *Working Papers 52, Lund University, Department of Linguistics and Phonetics*, 5–8.

Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge University Press.

Baldeweg, T. (2006). Repetition effects to sounds: Evidence for predictive coding in the auditory system. *Trends in Cognitive Sciences*, *10*(3), 93–94. https://doi.org/10.1016/j.tics.2006.01.010

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical Microcircuits for Predictive Coding. *Neuron*, *76*(4), 695–711. https://doi.org/10.1016/j.neuron.2012.10.038

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bendixen, A., Scharinger, M., Strauß, A., & Obleser, J. (2014). Prediction in the service of comprehension: Modulated early brain responses to omitted speech segments. *Cortex*, *53*, 9–26. https://doi.org/10.1016/j.cortex.2014.01.001

Boersma, P., & Weenink, D. (2022). *Praat: Doing phonetics by computer* (Version 6.2.09). http://www.praat.org/

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2019). Toward a Neurobiologically Plausible Model of Language-Related, Negative Event-Related Potentials. *Frontiers in Psychology*, *10*. https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00298

Boulenger, V., Hoen, M., Jacquier, C., & Meunier, F. (2011). Interplay between acoustic/phonetic and semantic processes during spoken sentence comprehension: An ERP study. *Brain and Language*, *116*(2), 51–63. https://doi.org/10.1016/j.bandl.2010.09.011

Brouwer, S., Mitterer, H., & Huettig, F. (2013). Discourse context and the recognition of reduced and canonical spoken words. *Applied Psycholinguistics*, *34*(3), 519–539. https://doi.org/10.1017/S0142716411000853

Brown, H., & Friston, K. J. (2012). Free-Energy and Illusions: The Cornsweet Effect. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00043

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

Bruce, G. (1987). How floating is focal accent? In K. Gregersen & H. Basbøll (Eds.), *Nordic prosody IV* (pp. 41–49). Odense University Press.

Bruce, G. (2007). Components of a prosodic typology of Swedish intonation. In T. Riad & C. Gussenhoven (Eds.), *Tones and Tunes: Typological Studies in Word and Sentence Prosody* (Vol. 1, pp. 113–146). Mouton de Gruyter. https://doi.org/10.1515/9783110207569.113

Brunellière, A., & Soto-Faraco, S. (2015). The interplay between semantic and phonological constraints during spoken-word comprehension. *Psychophysiology*, *52*(1), 46–58. https://doi.org/10.1111/psyp.12285

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/S0140525X12000477

Connolly, J. F., Byrne, J. M., & Dywan, C. A. (1995). Assessing adult receptive vocabulary with event-related potentials: An investigation of cross-modal and cross-form priming. *Journal of Clinical* *and Experimental Neuropsychology*, *17*(4), 548–565. https://doi.org/10.1080/01688639508405145

Connolly, J. F., & Phillips, N. A. (1994). Event-Related Potential Components Reflect Phonological and Semantic Processing of the Terminal Word of Spoken Sentences. *Journal of Cognitive Neuroscience*, *6*(3), 256–266. https://doi.org/10.1162/jocn.1994.6.3.256

Connolly, J. F., Phillips, N. A., Stewart, S. H., & Brake, W. G. (1992). Event-related potential sensitivity to acoustic and semantic properties of terminal words in sentences. *Brain and Language*, *43*(1), 1–18. https://doi.org/10.1016/0093-934X(92)90018-A

Connolly, J. F., Service, E., D'Arcy, R. C. N., Kujala, A., & Alho, K. (2001). Phonological aspects of word recognition as revealed by high-resolution spatio-temporal brain mapping: *Neuroreport*, *12*(2), 237–243. https://doi.org/10.1097/00001756-200102120-00012

D'Arcy, R. C. N., Connolly, J. F., & Crocker, S. F. (2000). Latency shifts in the N2b component track phonological deviations in spoken words. *Clinical Neurophysiology*, *111*(1), 40–44. https://doi.org/10.1016/S1388-2457(99)00210-2

Dehaene-Lambertz, G., Dupoux, E., & Gout, A. (2000). Electrophysiological Correlates of Phonological Processing: A Cross-linguistic Study. *Journal of Cognitive Neuroscience*, *12*(4), 635–647. https://doi.org/10.1162/089892900562390

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117–1121. https://doi.org/10.1038/nn1504

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*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Desroches, A. S., Newman, R. L., & Joanisse, M. F. (2009). Investigating the Time Course of Spoken Word Recognition: Electrophysiological Evidence for the Influences of Phonological Similarity. *Journal of Cognitive Neuroscience*, *21*(10), 1893–1906. https://doi.org/10.1162/jocn.2008.21142

Eckstein, K., & Friederici, A. D. (2006). It's Early: Event-related Potential Evidence for Initial Interaction of Syntax and Prosody in Speech Comprehension. *Journal of Cognitive Neuroscience*, *18*(10), 1696–1711. https://doi.org/10.1162/jocn.2006.18.10.1696

Elert, C.-C. (1972). Tonality in Swedish: Rules and a list of minimal pairs. In E. S. Firchow (Ed.), *Studies for Einar Haugen* (pp. 151–173). De Gruyter. https://doi.org/10.1515/9783110879131-015

Engstrand, O. (1995). Phonetic Interpretation of the Word Accent Contrast in Swedish. *Phonetica*, *52*(3), 171–179. https://doi.org/10.1159/000262168

Engstrand, O. (1997). Phonetic interpretation of the word accent contrast in Swedish: Evidence from spontaneous speech. *Phonetica*, *54*(2), 61–75.

Ernestus, M., Baayen, H., & Schreuder, R. (2002). The Recognition of Reduced Word Forms. *Brain and Language*, *81*(1–3), 162–173. https://doi.org/10.1006/brln.2001.2514

Felder, V., Jonsson-Steiner, E., Eulitz, C., & Lahiri, A. (2009). Asymmetric processing of lexical tonal contrast in Swedish. *Attention, Perception & Psychophysics*, *71*(8), 1890–1899. https://doi.org/10.3758/APP.71.8.1890

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for “top-down” effects. *Behavioral and Brain Sciences*, *39*, e229. https://doi.org/10.1017/S0140525X15000965

Friston, K., & Kiebel, S. (2009). Cortical circuits for perceptual inference. *Neural Networks*, *22*(8), 1093–1104. https://doi.org/10.1016/j.neunet.2009.07.023

Gårding, E., & Lindblad, P. (1973). Constancy and variation in Swedish word accent patterns. *Working Papers 7, Lund University, Department of Linguistics and Phonetics*.

Garellek, M. (2019). The phonetics of voice. In W. F. Katz & P. A. Assmann (Eds.), *The Routledge Handbook of Phonetics* (pp. 75–106). Routledge.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*(2), 251–279. https://doi.org/10.1037/0033-295X.105.2.251

Granqvist, S. (2009). *Sopran*. Tolvan Data. http://www.tolvan.com/index.php?page=/sopran/sopran.php

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, *48*(12), 1711–1725. https://doi.org/10.1111/j.1469-8986.2011.01273.x

Haenschel, C., Vernon, D. J., Dwivedi, P., Gruzelier, J. H., & Baldeweg, T. (2005). Event-Related Brain Potential Correlates of Human Auditory Sensory Memory-Trace Formation. *The Journal of Neuroscience*, *25*(45), 10494–10501. https://doi.org/10.1523/JNEUROSCI.1227-05.2005

Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech: Semantic ERP effects. *Neuropsychologia*, *38*(11), 1518–1530. https://doi.org/10.1016/S0028-3932(00)00052-X

Heldner, M. (2001). On the non-linear lengthening of focally accented Swedish words. *Nordic Prosody : Proceedings of the VIIIth Conference, Trondheim 2000*, 103–112.

Helmholtz, H. von. (1896). *Handbuch der physiologischen Optik* (2 umgearb. Auflage). L. Voss.

Homan, R. W., Herman, J., & Purdy, P. (1987). Cerebral location of international 10–20 system electrode placement. *Electroencephalography and Clinical Neurophysiology*, *66*(4), 376–382. https://doi.org/10.1016/0013-4694(87)90206-9

Hyman, L. M. (2009). How (not) to do phonological typology: The case of pitch-accent. *Language Sciences*, *31*(2–3), 213–238. https://doi.org/10.1016/j.langsci.2008.12.007

Ito, A., Martin, A. E., & Nieuwland, M. S. (2017). How robust are prediction effects in language comprehension? Failure to replicate article-elicited N400 effects. *Language, Cognition and Neuroscience*, *32*(8), 954–965. https://doi.org/10.1080/23273798.2016.1242761

Johnson, K. (1997). Speech Perception without Speaker Normalization. An exemplar model. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (Ed. 1, pp. 145–165). Academic Press.

Jun, S.-A. (Ed.). (2005). *Prosodic typology: The phonology of intonation and phrasing*. Oxford University Press.

Khoei, M. A., Masson, G. S., & Perrinet, L. U. (2017). The Flash-Lag Effect as a Motion-Based Predictive Shift. *PLOS Computational Biology*, *13*(1), e1005068. https://doi.org/10.1371/journal.pcbi.1005068

Kogo, N., & Trengove, C. (2015). Is predictive coding theory articulated enough to be testable? *Frontiers in Computational Neuroscience*, *9*. https://doi.org/10.3389/fncom.2015.00111

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, *55*(3–4), 243–276. https://doi.org/10.1556/ALing.55.2008.3-4.2

Kumar, S., Sedley, W., Nourski, K. V., Kawasaki, H., Oya, H., Patterson, R. D., Howard, M. A., Friston, K. J., & Griffiths, T. D. (2011). Predictive Coding and Pitch Processing in the Auditory Cortex. *Journal of Cognitive Neuroscience*, *23*(10), 3084–3094. https://doi.org/10.1162/jocn_a_00021

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kwon, J. (2023). *The role of word accents in semantic processing in South Swedish—An ERP study with minimal pairs* [Master's thesis, Lund University]. https://www.lu.se/lup/publication/9111821

Lewendon, J., Mortimore, L., & Egan, C. (2020). The Phonological Mapping (Mismatch) Negativity: History, Inconsistency, and Future Direction. *Frontiers in Psychology*, *11*, 1967. https://doi.org/10.3389/fpsyg.2020.01967

Luck, S. J. (2014). *An introduction to the event-related potential technique* (Second edition). The MIT Press.

Lupyan, G., & Clark, A. (2015). Words and the World: Predictive Coding and the Language-Perception-Cognition Interface. *Current Directions in Psychological Science*, *24*(4), 279–284. https://doi.org/10.1177/0963721415570732

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1–2), 71–102. https://doi.org/10.1016/0010-0277(87)90005-9

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

Mitterer, H., & Blomert, L. (2003). Coping with phonological assimilation in speech perception: Evidence for early compensation. *Perception & Psychophysics*, *65*(6), 956–969. https://doi.org/10.3758/BF03194826

Myrberg, S. (2010). *The intonational phonology of Stockholm Swedish* [Doctoral Dissertation]. Acta Univ. Stockholmiensis.

Myrberg, S., & Riad, T. (2015). The prosodic hierarchy of Swedish. *Nordic Journal of Linguistics*, *38*(2), 115–147. https://doi.org/10.1017/S0332586515000177

Näätänen, R., Kujala, T., Escera, C., Baldeweg, T., Kreegipuu, K., Carlson, S., & Ponton, C. (2012). The mismatch negativity (MMN) – A unique window to disturbed central auditory processing in ageing and different clinical conditions. *Clinical Neurophysiology*, *123*(3), 424–458. https://doi.org/10.1016/j.clinph.2011.09.020

Näätänen, R., Paavilainen, P., Titinen, H., Jiang, D., & Alho, K. (1993). Attention and mismatch negativity. *Psychophysiology*, *30*(5), 436–450. https://doi.org/10.1111/j.1469-8986.1993.tb02067.x

Newman, R. L., & Connolly, J. F. (2009). Electrophysiological markers of pre-lexical speech processing: Evidence for bottom–up and top–down effects on spoken word processing. *Biological Psychology*, *80*(1), 114–121. https://doi.org/10.1016/j.biopsycho.2008.04.008

Newman, R. L., Connolly, J. F., Service, E., & Mcivor, K. (2003). Influence of phonological expectations during a phoneme deletion task: Evidence from event-related brain potentials. *Psychophysiology*, *40*(4), 640–647. https://doi.org/10.1111/1469-8986.00065

Nicenboim, B., Vasishth, S., & Rösler, F. (2020). Are words pre-activated probabilistically during sentence comprehension? Evidence from new data and a Bayesian random-effects meta-analysis using publicly available data. *Neuropsychologia*, *142*, 107427. https://doi.org/10.1016/j.neuropsychologia.2020.107427

Nieuwland, M. S. (2019). Do 'early' brain responses reveal word form prediction during language comprehension? A critical review. *Neuroscience & Biobehavioral Reviews*, *96*, 367–400. https://doi.org/10.1016/j.neubiorev.2018.11.019

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Von Grebmer Zu Wolfsthurn, S., Bartolozzi, F., Kogan, V., Ito, A., Mézière, D., Barr, D. J., Rousselet, G. A., Ferguson, H. J., Busch-Moreno, S., Fu, X., Tuomainen, J., Kulakova, E., Husband, E. M., … Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *ELife*, *7*, e33468. https://doi.org/10.7554/eLife.33468

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*(3), 189–234. https://doi.org/10.1016/0010-0277(94)90043-4

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, *2011*, 1–9. https://doi.org/10.1155/2011/156869

Otten, M., & Van Berkum, J. J. A. (2008). Discourse-Based Word Anticipation During Language Processing: Prediction or Priming? *Discourse Processes*, *45*(6), 464–496. https://doi.org/10.1080/01638530802356463

Petkov, C. I., & Sutter, M. L. (2011). Evolutionary conservation and neuronal mechanisms of auditory perceptual restoration. *Hearing Research*, *271*(1–2), 54–65. https://doi.org/10.1016/j.heares.2010.05.011

Port, R. F., & Leary, A. P. (2005). Against Formal Phonology. *Language*, *81*(4), 927–964. https://doi.org/10.1353/lan.2005.0195

Poulton, V. R., & Nieuwland, M. S. (2022). Can You Hear What's Coming? Failure to Replicate ERP Evidence for Phonological Prediction. *Neurobiology of Language*, *3*(4), 556–574. https://doi.org/10.1162/nol_a_00078

Riad, T. (2013). *The Phonology of Swedish*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199543571.001.0001

Riecke, L., Esposito, F., Bonte, M., & Formisano, E. (2009). Hearing Illusory Sounds in Noise: The Timing of Sensory-Perceptual Transformations in Auditory Cortex. *Neuron*, *64*(4), 550–561. https://doi.org/10.1016/j.neuron.2009.10.016

Roll, M. (2015). A neurolinguistic study of South Swedish word accents: Electrical brain potentials in nouns and verbs. *Nordic Journal of Linguistics*, *38*(2), 149–162. https://doi.org/10.1017/S0332586515000189

Roll, M. (2022). The predictive function of Swedish word accents. *Frontiers in Psychology*, *13*(910787), 1–11. https://doi.org/10.3389/fpsyg.2022.910787

Roll, M., & Horne, M. (2011). Interaction of right- and left-edge prosodic boundaries in syntactic parsing. *Brain Research*, *1402*, 93–100. https://doi.org/10.1016/j.brainres.2011.06.002

Roll, M., Horne, M., & Lindgren, M. (2010). Word accents and morphology—ERPs of Swedish word processing. *Brain Research*, *1330*, 114–123. https://doi.org/10.1016/j.brainres.2010.03.020

Roll, M., Söderström, P., & Horne, M. (2013). Word-stem tones cue suffixes in the brain. *Brain Research*, *1520*, 116–120. https://doi.org/10.1016/j.brainres.2013.05.013

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. https://doi.org/10.1016/j.bandl.2015.07.009

RStudio Team. (2022). *RStudio: Integrated Development Environment for R*. PBC. http://www.rstudio.com/

Samuel, A. G. (1996). Does lexical information influence the perceptual restoration of phonemes? *Journal of Experimental Psychology: General*, *125*(1), 28–51. https://doi.org/10.1037/0096-3445.125.1.28

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language: Music and prosody: An ERP study. *Psychophysiology*, *41*(3), 341–349. https://doi.org/10.1111/1469-8986.00172.x

Söderström, P., Horne, M., & Roll, M. (2017). Stem Tones Pre-activate Suffixes in the Brain. *Journal of Psycholinguistic Research*, *46*(2), 271–280. https://doi.org/10.1007/s10936-016-9434-2

Svensson Lundmark, M., Ambrazaitis, G., & Ewald, O. (2017). Exploring Multidimensionality: Acoustic and Articulatory Correlates of Swedish Word Accents. *Interspeech 2017*, 3236–3240. https://doi.org/10.21437/Interspeech.2017-1502

Svensson Lundmark, M., Frid, J., & Schötz, S. (2015). A pilot study: Acoustic and articulatory data on tonal alignment in Swedish word accents. *Working Papers 55, Lund University, Department of Linguistics and Phonetics*, 75–80.

van den Brink, D., Brown, C. M., & Hagoort, P. (2001). Electrophysiological Evidence for Early Contextual Influences during Spoken-Word Recognition: N200 Versus N400 Effects. *Journal of Cognitive Neuroscience*, *13*(7), 967–985. https://doi.org/10.1162/089892901753165872

van den Brink, D., & Hagoort, P. (2004). The Influence of Semantic and Syntactic Context Constraints on Lexical Selection and Integration in Spoken-Word Comprehension as Revealed by ERPs. *Journal of Cognitive Neuroscience*, *16*(6), 1068–1084. https://doi.org/10.1162/0898929041502670

Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 394–417. https://doi.org/10.1037/0278-7393.25.2.394

Wacongne, C., Labyt, E., Van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences*, *108*(51), 20754–20759. https://doi.org/10.1073/pnas.1117807108

Winter, B. (2020). *Statistics for linguists: An introduction using R*. Routledge, Taylor & Francis Group.

Zwitserlood, P. (2004). Sublexical and morphological information in speech processing. *Brain and Language*, *90*(1–3), 368–377. https://doi.org/10.1016/S0093-934X(03)00448-6

Appendix

Appendix A1: Flyer used during the recruitment phase

PARTICIPATE IN A LANGUAGE STUDY AND GET A MOVIE TICKET!

We're conducting an EEG (electroencephalography) study on the way Swedish speakers comprehend sentences and looking for participants who meet the following criteria:

- between 18 and 40 years old
- right-handed
- only Swedish as the native language
- grown up in STOCKHOLM (it's fine if you live in Skåne now)
- no hearing impairment

If you fit these criteria and want to participate, please contact Jule at the following email-address:

ju2041na-s@student.lu.se

As a thank you for your participation, you will get a voucher for a movie ticket of your choice at Filmstaden.

Appendix A2: Text used during the recruitement phase

“EEG-study on sentence perception - native speakers of Stockholm Swedish wanted!

Hi! I'm currently looking for participants for a study that I'm doing for my MA-thesis in the program "Language and Linguistics". The study will be conducted at the Humanities Lab at Lund University. It is an EEG (electroencephalographic) study where we measure brain waves by placing a cap with electrodes on the participants head.

The study takes about 2 hours and pays 100kr in form of a movie ticket to Filmstaden. After applying the cap, you will listen to sentences and occasionally answer comprehension questions about the sentences you heard.

Participants should be 18-40 years old, right-handed and should originally be from STOCKHOLM and have ONLY Swedish as a first language. This means that you grew up in Stockholm or the surrounding areas. It's fine if you live in Lund now. You're welcome to contact me if you have questions about this or any other parts of the study.

Looking forward to working with you :)”

Appendix B: Sentences used as stimulus material

Condition	Context sentence	Constant sentence	Target word
arg_pl_hi	Barnen var inte glada utan	de var	arga
arg_pl_lo	Politikerna såg ut som	de var	arga
arg_sg_hi	Vi hade förväntat oss en glad granne men	han var	arg
arg_sg_lo	Jag pratade inte med honom därför att	han var	arg
dum_pl_hi	Han förolämpade deras intelligens när han sa att	de var	dumma

dum_pl_lo	Om jag skulla vara ärlig tyckte jag att	de var	dumma
dum_sg_hi	Hunden var inte smart utan	den var	dum
dum_sg_lo	När man tittade på deras katt kunde man se att	den var	dum
dyr_pl_hi	Jag var tvungen att spara för att köpa skorna för	de var	dyra
dyr_pl_lo	Alla kände till bilderna för att	de var	dyra
dyr_sg_hi	Hon sa att osten var billig men	den var	dyr
dyr_sg_lo	Mangon var enkel att få tag i men	den var	dyr
feg_pl_hi	Vi antog att lejonen skulle vara modiga men	de var	fega
feg_pl_lo	Hon hade hört att	de var	fega
feg_sg_hi	Han var inte modig utan	han var	feg
feg_sg_lo	Vad jag inte visste om mannen var att	han var	feg
fin_pl_hi	Bröden var inte grova utan	de var	fina
fin_pl_lo	Han visste inte mycket om blommor förutom att	de var	fina
fin_sg_hi	Doften luktade inte illa utan	den var	fin
fin_sg_lo	När han frågade om tavlans egenskaper mindes jag bara att	den var	fin
ful_pl_hi	De nya husen var inte fina, utan	de var	fula
ful_pl_lo	Böckerna såldes inte för att	de var	fula
ful_sg_hi	Han tyckte bilden var fin, men jag tyckte	den var	ful
ful_sg_lo	Efter att han såg vasen så sa han att han tyckte att	den var	ful
glad_pl_hi	Föräldrarna var inte ledsna utan	de var	glada
glad_pl_lo	Tjejerna kunde inte visa att	de var	glada
glad_sg_hi	Vi förväntade oss en ledsen kvinna men	hon var	glad
glad_sg_lo	Det var en bra dag för	hon var	glad
grön_pl_hi	Kvinnan tyckte om höstens färger men hon föredrog träden när	de var	gröna
grön_pl_lo	Det var enkelt att hitta böckerna i bokhyllan för	de var	gröna
grön_sg_hi	Gräsmattan var inte gul utan	den var	grön
grön_sg_lo	Eftersom jag inte sett soffan visste jag inte att	den var	grön
grov_pl_hi	Brotten var inte bara vårdslösa utan	de var	grova
grov_pl_lo	Han kasserade linorna för	de var	grova
grov_sg_hi	Brödlimpan var inte fin utan	den var	grov
grov_sg_lo	Hon kritiserade indelningen för	den var	grov
grund_pl_hi	Älvarna var inte djupa utan	de var	grunda
grund_pl_lo	Kusthabitaten passade för fåglarna för att	de var	grunda
grund_sg_hi	Man kunde stå i floden för att	den var	grund
grund_sg_lo	När man tittade på sjön kunde man se att	den var	grund

gul_pl_hi	Bananerna var inte gröna utan	de var	gula
gul_pl_lo	Vi kunde inte använda färgerna för att	de var	gula
gul_sg_hi	Han påstod att solen var vit även om vi sa att	den var	gul
gul_sg_lo	Tjejen tyckte om tröjan för att	den var	gul
hal_pl_hi	Ålarna var inte sträva utan	de var	hala
hal_pl_lo	Han sa till oss att vi skulle se upp med golven i museet för	de var	hala
hal_sg_hi	På grund av ishalkan kunde man inte köra på gatan för	den var	hal
hal_sg_lo	Vägen var avstängd för	den var	hal
hård_pl_hi	Det var inte bekvämt att sitta på stolarna för	de var	hårda
hård_pl_lo	Färgerna passade inte in i rummet för att	de var	hårda
hård_sg_hi	Madrassen var inte mjuk utan	den var	hård
hård_sg_lo	Jag ville inte köpa soffan för	den var	hård
hög_pl_hi	Bergen var inte låga utan	de var	höga
hög_pl_lo	Han ville köpa husen för	de var	höga
hög_sg_hi	Man kunde se flaggstången på långt avstand för att	den var	hög
hög_sg_lo	Jag tyckte hatten var ful för	den var	hög
kall_pl_hi	Bullarna var inte varma utan	de var	kalla
kall_pl_lo	När hon hämtade skedarna märkte hon att	de var	kalla
kall_sg_hi	Jag var tvungen att värma upp maträtten för	den var	kall
kall_sg_lo	Det gick inte att starta bilen för	den var	kall
känd_pl_hi	Sångarna i gruppen var inte okända längre utan	de var	kända
känd_pl_lo	Jag valde ut sångerna för	de var	kända
känd_sg_hi	Fotograferna följde skådespelaren för	han var	känd
känd_sg_lo	Alla brydde sig om historien för	den var	känd
låg_pl_hi	Temperaturerna var inte höga utan	de var	låga
låg_pl_lo	Skåpen såg bra ut i rummet för	de var	låga
låg_sg_hi	Min hjärtfrekvens var bara 40 så doktorn sa att	den var	låg
låg_sg_lo	Jag ville inte höra rösten för	den var	låg
lång_pl_hi	Han sa att tyckte mest om giraffers halsar för att	de var	långa
lång_pl_lo	Hela familjen var känd för att	de var	långa
lång_sg_hi	Pojken var inte kort längre utan	han var	lång
lång_sg_lo	Mannen var jättestolt över att	han var	lång
lugn_pl_hi	Gästerna var inte längre upprörda utan	de var	lugna
lugn_pl_lo	När vi arbetade med hästerna var det viktigt att	de var	lugna
lugn_sg_hi	Sjön var inte stormig utan	den var	lugn
lugn_sg_lo	Platsen var mitt favoritställe i stan för	den var	lugn

ny_pl_hi	Efter städningen såg de gamla mattorna ut som om	de var	nya
ny_pl_lo	Folk var nyfikna på butikerna i köpcentret för	de var	nya
ny_sg_hi	Klänningen var inte gammal utan	den var	ny
ny_sg_lo	Skolan hade inte så många elever för att	den var	ny
rå_pl_hi	Biffarna såg ganska röda ut för	de var	råa
rå_pl_lo	Grönsakerna såg ut som om	de var	råa
rå_sg_hi	Köttbullen var inte stekt utan	den var	rå
rå_sg_lo	Pojken tyckte inte om potatisen för	den var	rå
ren_pl_hi	Jag förväntade mig att skorna skulle vara smutsiga, men	de var	rena
ren_pl_lo	Cyklarna såg fina ut för	de var	rena
ren_sg_hi	Mattan var inte smutsig utan	den var	ren
ren_sg_lo	Han var mycket glad över duschen för	den var	ren
röd_pl_hi	Äpplena var inte gröna utan	de var	röda
röd_pl_lo	Mattorna passade inte till inredningen för	de var	röda
röd_sg_hi	Hon kunde inte se rosens färg så jag berättade att	den var	röd
röd_sg_lo	Jag använde pennan för att	den var	röd
sann_pl_hi	Hennes påståenden var inte falska utan	de var	sanna
sann_pl_lo	Det som förvånade mig mest om historierna var att	de var	sanna
sann_sg_hi	Han hade många bevis för sin historia så jag insåg att	den var	sann
sann_sg_lo	Han påstod att	den var	sann
sen_pl_hi	De hann inte till tidpunkten och skrev till oss att	de var	sena
sen_pl_lo	De tyckte jättemycket om föreställningen även om	de var	sena
sen_sg_hi	Bussen var inte tidig utan	den var	sen
sen_sg_lo	Problemet med posten var att	den var	sen
smal_pl_hi	Älvarna var inte breda utan	de var	smala
smal_pl_lo	Många tyckte inte om de nya fönstren för	de var	smala
smal_sg_hi	Vi förväntade oss en bred cykelfil men	den var	smal
smal_sg_lo	Bron var känd för att	den var	smal
snabb_pl_hi	Till min stora förvåning var sniglarna inte långsamma utan	de var	snabba
snabb_pl_lo	Hon tyckte om kaninerna för att	de var	snabba
snabb_sg_hi	Bilen var inte långsam utan	den var	snabb
snabb_sg_lo	Vi kunde inte se bilen men	den var	snabb
snål_pl_hi	De ville inte ge pengar till oss för	de var	snåla
snål_pl_lo	Det fanns en anledning till att	de var	snåla
snål_sg_hi	Frun var inte generös utan	hon var	snål
snål_sg_lo	Kvinnan fick dåligt rykte för	hon var	snål

stängd_pl_hi	Butikerna var inte öppna utan	de var	stängda
stängd_pl_lo	Föräldrarna kritiserade skolorna för	de var	stängda
stängd_sg_hi	Vi kunde inte gå genom dörren för	den var	stängd
stängd_sg_lo	Kafeterian sålde inga kakor för	den var	stängd
stor_pl_hi	Skeppen var inte små utan	de var	stora
stor_pl_lo	Hon visste inte hur katterna såg ut men hon visste att	de var	stora
stor_sg_hi	Jag mindes katten som när den var liten, men nu såg jag att	den var	stor
stor_sg_lo	Konstnären var stolt över tavlan för att	den var	stor
svag_pl_hi	Det var ganska svårt att känna dofterna så jag sa att	de var	svaga
svag_pl_lo	Alla patienter fick soppa för att	de var	svaga
svag_sg_hi	Armen var inte stark utan	den var	svag
svag_sg_lo	Hunden stannade hemma för	den var	svag
svår_pl_hi	Läxorna var inte enkla utan	de var	svåra
svår_pl_lo	Hon markerade uppgifterna för att	de var	svåra
svår_sg_hi	Han sa att gåtan var enkel men jag tyckte	den var	svår
svår_sg_lo	Hon sjöng inte låten därför att	den var	svår
tam_pl_hi	Hundarna såg vilda ut även om husse sa att	de var	tama
tam_pl_lo	Han kritiserade inläggen i debatten för	de var	tama
tam_sg_hi	Då vi adopterade katten var den inte längre vild utan	den var	tam
tam_sg_lo	Alla älskade björnen därför att	den var	tam
tom_pl_hi	Jag trodde flaskorna var fulla, men	de var	tomma
tom_pl_lo	Han tyckte inte om rummen för	de var	tomma
tom_sg_hi	Muggen var inte full utan	den var	tom
tom_sg_lo	Jag använde lådan för	den var	tom
torr_pl_hi	Jag torktumlade handdukarna tills	de var	torra
torr_pl_lo	Jag tog hand om växterna för	de var	torra
torr_sg_hi	Öknen var inte våt utan	den var	torr
torr_sg_lo	Clementinen såg ut som om	den var	torr
trång_pl_hi	Tunnlarna var inte vida utan	de var	trånga
trång_pl_lo	Han kritiserade tågkupéerna för	de var	trånga
trång_sg_hi	Hon hatade att ta hissen för	den var	trång
trång_sg_lo	Hon tyckte inte om cykelbanan för	den var	trång
tung_pl_hi	Vikterna var inte lätta,	de var	tunga

tung_pl_lo	Vad han inte sa om dörrarna var att	de var	tunga
tung_sg_hi	Jag antog att resväskan var lätt men	den var	tung
tung_sg_lo	Jag tog inte med mig boken för	den var	tung
tunn_pl_hi	Såserna var inte tjocka utan	de var	tunna
tunn_pl_lo	De kunde inte använda jackorna för	de var	tunna
tunn_sg_hi	Vi fick inte gå på isen för	den var	tunn
tunn_sg_lo	Jag kunde inte riktig se personen men jag såg att	den var	tunn
ung_pl_hi	Lärarna var inte gamla utan	de var	unga
ung_pl_lo	Fåglarna var väldigt söta för	de var	unga
ung_sg_hi	De trodde att kvinnan var gammal, men	hon var	ung
ung_sg_lo	Jag visste inte att	hon var	ung
varm_pl_hi	Kläderna var inte kalla utan	de var	varma
varm_pl_lo	Jag minndes de dagarna väl för	de var	varma
varm_sg_hi	Glassen skulle smälta om	den var	varm
varm_sg_lo	Lampan såg ut som	den var	varm

Appendix C1: Biographical questionnaire

Biographical Data Questionnaire
for a study
on sentence comprehension

Participant code	
Age	
Native language	
Where did you spend most of your life? (Country, region)	
Do you speak any other languages? Please name them.	
Do you have a hearing impairment that hasn't been corrected?	
Have you been diagnosed with a neurological condition? (e.g. Dementia, Parkinson, …)	Yes No Other (please specify):
Which is your dominant hand?	

Appendix C2: Consent form

LUNDS UNIVERSITET

Informed consent to participate in study
on sentence comprehension

1.	**Background and purpose**	This study is part of a Master's degree thesis at the department for Languages and Literature (SOL), at Lund University, with Mechtild Tronnier and Mikael Roll as supervisors. The purpose is to investigate the time course of sentence comprehension in spoken language.
2.	**The study**	For this purpose, you will listen two 160 spoken sentences and answer comprehension questions at different intervals. To measure brain activity, an EEG cap will be placed on your head and a conducive gel is applied between the electrode and the scalp.
3.	**Handling and storing the data**	All data is anonymized. We do not collect any information that would be sufficient to identify you later. Observe that the supervisors will have access to the raw data. The processed version of the anonymized EEG data from all participants, as well as summaries of biographical information about all participants (age, gender) will be published together with the Master's thesis or other follow-up publications, but not forwarded to third parties. *The long-term storage of the data will follow the Lund university policy for data storage.*
4.	**Voluntary participation**	Participation is voluntary, and as a participant you have the right to cease participation at any time. You also have the right to request the deletion of your data at any point to the extent that this is possible once the results from the experiment have been published.
5.	**Responsible persons**	**Jule Nabrotzky** **+46[redacted]** **ju2041na-s@student.lu.se** **Supervisors: Mechtild Tronnier, Mikael Roll** **mechtild.tronnier@ling.lu.se, mikael.roll@ling.lu.se**

I confirm by my signature that I have read the information about the study, and that I consent to participate. This form is made in two copies, one for me, and one for the responsible person.

____
Date, Place

____
Signature / clarification of signature /

Appendix D: Comprehension questions

Sentence	question	Correct answer
Vi hade förväntat oss en glad granne men han var arg	Var det läraren som var arg?	nej
Barnen var inte glada utan de var arga	Var det tjejerna som var arga?	nej
Hunden var inte smart utan den var dum	Var det katten som var dum?	nej
Mangon var enkel att få tag i men den var dyr	Var det mangon som var dyr?	ja
Jag var tvungen att spara för att köpa skorna för de var dyra	Var det tröjorna som var dyra?	nej
Vad jag inte visste om mannen var att han var feg	Var det mannen som var feg?	ja
Vi antog att lejonen skulle vara modiga men de var fega	Var det lejonen som var fega?	ja
När han frågade om tavlans egenskaper mindes jag bara att den var fin	Var det bilden som var fin?	nej
Han visste inte mycket om blommor förutom att de var fina	Var det blommor som var fina?	ja
Bröden var inte grova utan de var fina	Var det bröden som var grova?	ja
Han tyckte bilden var fin, men jag tyckte den var ful	Var det bilden som var ful?	ja
Böckerna såldes inte för att de var fula	Var det magasinen som var fula?	nej
Vi förväntade oss en ledsen kvinna men hon var glad	Var det kvinnan som var glad?	ja
Tjejerna kunde inte visa att de var glada	Var det kvinnorna som var glada?	nej
Eftersom jag inte sett soffan visste jag inte att den var grön	Var det stolen som var grön?	nej
Gräsmattan var inte gul utan den var grön	Var det lövet som var grönt?	nej
Det var enkelt att hitta böckerna i bokhyllan för de var gröna	Var det bokhyllorna som var gröna?	nej
Han påstod att solen var vit även om vi sa att den var gul	Var det månen som var gul?	nej

Bananerna var inte gröna utan de var gula	Var det äpplena som var gula?	nej
Vi kunde inte använda färgerna för att de var gula	Var det pennorna som var gula?	nej
Madrassen var inte mjuk utan den var hård	Var det madrassen som var hård?	ja
Jag ville inte köpa soffan för den var hård	Var det soffan som var hård?	ja
Färgerna passade inte in i rummet för att de var hårda	Var det färgerna som var hårda?	ja
Man kunde se flaggstången på långt avstand för att den var hög	Var det flaggstången som var hög?	ja
Bergen var inte låga utan de var höga	Var det tornen som var höga?	nej
Det gick inte att starta bilen för den var kall	Var det bilen som var kall?	ja
När hon hämtade skedarna märkte hon att de var kalla	Var det skedarna som var kalla?	ja
Bullarna var inte varma utan de var kalla	Var det bullarna som var kalla?	ja
Fotograferna följde skådespelaren för han var känd	Var det reportern som följde skådespelaren?	nej
Jag ville inte höra rösten för den var låg	Var det rösten som var låg?	ja
Min hjärtfrekvens var bara 40 så doktorn sa att den var låg	Var det mamman som jag pratade med?	nej
Skåpen såg bra ut i rummet för de var låga	Var det sängarna som såg bra ut?	nej
Mannen var jättestolt över att han var lång	Var det pojken som var lång?	nej
Han sa att tyckte mest om giraffers halsar för att de var långa	Var det benen som var långa?	nej
Sjön var inte stormig utan den var lugn	Var det havet som var lugnt?	nej
Gästerna var inte längre upprörda utan de var lugna	Var det gästerna som var lugna?	ja
Klänningen var inte gammal utan den var ny	Var det klänningen som var ny?	ja

Folk var nyfikna på butikerna i köpcentret för de var nya	Var det butikerna som var nya?	ja
Mattan var inte smutsig utan den var ren	Var det kudden som var ren?	nej
Han var mycket glad över duschen för den var ren	Var det duschen som var ren?	ja
Cyklarna såg fina ut för de var rena	Var det bilarna som såg fina ut?	nej
Jag använde pennan för att den var röd	Var det boken som jag använde?	nej
Hon kunde inte se rosens färg så jag berättade att den var röd	Var det tulpanen som var röd?	ja
Mattorna passade inte till inredningen för de var röda	Var det mattorna som var röda?	ja
Bussen var inte tidig utan den var sen	Var det bussen som var sen?	ja
De tyckte jättemycket om föreställningen även om de var sena	Tyckte de om föreställningen?	ja
Bron var känd för att den var smal	Var det bron som var smal?	ja
Många tyckte inte om de nya fönstren för de var smala	Var det fönstren som var smala?	ja
Vi kunde inte se bilen men den var snabb	Var det cykeln som var snabb?	nej
Kafeterian sålde inga kakor för den var stängd	Var det restaurangen som sålde kakor?	nej
Vi kunde inte gå genom dörren för den var stängd	Var det ingången som var stängd?	ja
Skeppen var inte små utan de var stora	Var det skeppen som var stora	ja
Hon visste inte hur katterna såg ut men hon visste att de var stora	Var det hundarna som var stora?	nej
Alla patienter fick soppa för att de var svaga	Var det soppa som patienterna fick?	ja
Då vi adopterade katten var den inte längre vild utan den var tam	Var det hunden som var tam?	nej
Jag trodde flaskorna var fulla, men de var tomma	Var det flaskorna som var fulla?	nej
Clementinen såg ut som om den var torr	Var det äpplet som såg torr ut?	nej

Jag torktumlade handdukarna tills de var torra	Var det strumporna som jag torktumlade?	nej
De trodde att kvinnan var gammal, men hon var ung	Var det tanten som var gammal?	nej
Fåglarna var väldigt söta för de var unga	Var det fåglarna som var söta?	ja
Kläderna var inte kalla utan de var varma	Var det kläderna som var varma?	ja