Master's thesis

2023

Viktoriia Afoian

Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Humanities
Department of Language and Literature

NTNU
Norwegian University of
Science and Technology

Viktoriia Afoian

The Online Processing of Grammatical Gender and Number Agreeement in Norwegian

An ERP study

November 2023

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

The Online Processing of Grammatical Gender and Number Agreeement in Norwegian

An ERP study

Viktoriia Afoian

English Linguistics and Language Acquisition
Submission date: November 2023
Supervisor: Giosuè Baggio
Co-supervisor: Mila Vulchanova

Norwegian University of Science and Technology
Department of Language and Literature

Abstract

While grammatical gender is often described as a lexically stored feature that lacks a direct semantic interpretation, grammatical number has a referential role, providing information about the plurality of the entity to which it refers. Existing literature, however, is inconclusive as to whether these inherent representational differences affect the processing of gender and number agreement relations.

In the present work, we use event-related potentials (ERPs) to examine how morphosyntactic agreement is processed at the brain level. Employing the mismatch paradigm, we manipulate number and gender agreement in short sentences: the adjective either agrees or disagrees with the noun in terms of gender or number. Participants read Norwegian sentences, while the electroencephalogram (EGG) was recorded from 32 electrodes. The results did not yield significant effects for either the gender or number conditions, making it challenging to draw definitive conclusions regarding differences between the features. However, even though the results did not reach significance level, the effects elicited by gender and number violations appear to differ qualitatively.

We further explore the online processing of agreement relations in sentences with pseudowords, particularly if such relations can be computed when the meaning of a word cannot be accessed. We make use of pseudowords that follow the phonological and morphological patterns of Norwegian but lack lexical representations. In fact, the results revealed the effects of agreement with real words, but not with pseudowords. This implies agreement relations cannot be established just between abstract features and do require words with lexical representation.

Finally, we compare real words and pseudowords to explore the effects of composition. This comparison showed a P600 effect, with real words yielding a more positive response than pseudowords. At the same time, no effects were found with incorrect sentences. These results suggest that the P600 component might be associated with semantic and syntactic composition, and that the presence of morphosyntactic errors can inhibit compositional processes.

Abstrakt

Grammatisk kjønn (genus) blir ofte beskrevet som en leksikalsk kategori som mangler en direkte semantisk betydning. Derimot har grammatisk tall (numerus) en refererende rolle, som gir tydelig informasjon om mengden enheter det refereres til (omtales). Eksisterende litteratur er mangelful når det gjelder hvordan disse iboende forskjellene i representasjon påvirker prosessering av grammatisk samsvarsbøyning i kjønn og tall.

I denne masteroppgaven benytter vi hendelsesrelaterte potensialer (ERPs) for å undersøke hvordan morfosyntaktisk kongruensbøyning prosesseres i hjernen. Ved å benytte et violation-paradigme manipulerer vi samsvarsbøyning i kjønn og tall i korte setninger: adjektivet blir enten samsvarsbøyd eller ikke med substantivet i kjønn eller tall. Deltakerne leste norske setninger på PC-skjermen mens elektroencefalografi (EEG) ble registrert fra 32 elektroder. Resultatene viser ingen signifikante effekter for verken kjønn eller tall. Det gjør det utfordrende å trekke klare konklusjoner angående forskjeller mellom disse grammatiske kategoriene. Selv om resultatene ikke var signifikante, er det indikasjon på at feil i kjønns- og tallsamsvarsbøyning utløser effekter som er kvalitativt forskjellige.

Vi utforsker også online-prosessering av samsvarsbøyning, med et spesielt fokus på om slike relasjoner kan etableres når ord har ingen betydning. For dette bruker vi pseudord som følger norske fonologiske og morfologiske mønstre, men mangler leksikalske representasjoner. Resultatene viser effektene av samsvarsbøyning med ekte ord, men ikke med pseudord. Dette indikerer at samsvarsbøyningsrelasjoner ikke kan etableres bare mellom abstrakte kategorier og krever at elementer har leksikalske representasjoner.

Til slutt sammenligner vi ekte ord og pseudord for å utforske effekter av komposisjon. Denne analysen avslørte en P600-effekt, hvor ekte ord produserte en sterkere positiv respons enn pseudord. Imidlertid ble det ikke oppdaget noen effekter i setninger med feil. Dette antyder at P600 kan relateres til semantisk og grammatisk komposisjon, og at morfosyntaktiske feil kan hindre slike komposisjonsprosesser.

Acknowledgements

Completing this thesis has been a journey that I could not have navigated alone. The support, guidance, and encouragement I have received along the way have been invaluable.

First of all, I would like to thank my supervisors, Professor Giosuè Baggio and Professor Mila Vulchanova for their guidance throughout the process. A special thanks to Professor Giosuè Baggio for his time, for always being available for discussions, no matter the topic, and particularly for his assistance in data analysis.

I would also like to thank all the participants for taking part in the experiment. Without you, this thesis would have not been possible.

Lastly, but most importantly, I'm incredibly grateful for the support I received from Christoph Halser, Alevtina Roshchina, Julia Senderska, and Guro Sjuls. Your help throughout this journey means a lot to me.

Table of Contents

List of Figures

List of Tables

1 Introduction

Language processing unfolds at an extremely fast rate. Therefore, one central question of linguistic research is how humans are able to effectively and precisely handle complex linguistic structures. Researchers seek to understand the cognitive and neural mechanisms that underlie various aspects of language processing. This includes morphosyntactic processing, which involves operations related to the recognition and interpretation of grammatical structures and word forms. One approach to understand this phenomenon is to investigate grammatical features of a language, such as gender and number. Grammatical gender and number are the core properties of the noun and can be found in many languages. By examining gender and number agreement, valuable insights can be derived about the way the human brain stores both grammatical and lexical information. The grammatical categories of number and gender are often contrasted with each other in terms of their interpretability. Pesetsky and Torrego (2007) claim that the distinguishing factor between interpretable and uninterpretable features is whether a feature of a lexical item contributes to its semantic interpretation or not, allowing it to be understood in a meaningful way. Uninterpretable features do not have a direct semantic interpretation and serve as a tool for grammatical structure-building, allowing speakers to produce and comprehend grammatically correct sentences. In simple terms, the category of number can be described as referential, i.e. providing information about the plurality of the entity. In contrast, grammatical gender can be labelled as a lexically stored feature, considering it cannot be inferred solely based on the word's form or context. Assuming that the two features are different in the level of representation, the computation of agreement relations might involve distinct processing routines.

There is an ongoing debate about whether gender and number are processed and represented differently. Faussart et al. (1999) argue that gender is processed at an earlier stage than number due to its lexical nature. When processing gender mismatch, the parser needs to go back to the noun stem to access information about the noun and its grammatical gender. However, when there is a number agreement mismatch, the parser might be able to fix the mistake already at the syntactic level. Thereby, the processing of gender mismatch might employ more cognitive resources for analyzing and repairing the sentence. Conversely, other scholars support the idea that the processing of number mismatches is costlier due to the presence of multiple possible options for correction or reanalysis (Popov, 2022). Specifically, there can be only one possible option for gender (e.g. *et_{neuter} $stor_{common}$ hus_{neuter} vs et_{neuter} $stort_{neuter}$ hus_{neuter} 'a big house'), as compared to the number mismatch with two possible options for reanalysis (e.g. *et_{sg} $store_{pl}$ hus_{sg} vs et_{sg} $stort_{neut,sg}$ hus_{sg}/$store_{pl}$ hus_{pl} 'a big house/big houses'). The literature is inconclusive on what is more difficult to process: gender or number disagreement. While some studies indicate that gender disagreement is more demanding to process than number disagreement (Barber & Carreiras, 2005; Faussart et al., 1999; O'Rourke & Van Petten, 2011), others suggest the opposite or report no significant differences (Alemán Bañón, 2010; De Vincenzi, 1999; Dowens et al., 2010; Lukatela et al., 1987; Nevins et al., 2007). However, this inconclusiveness may be attributed to differences in experimental designs, participants, and stimuli employed in each study, as well as to the specific language under investigation. Therefore, more research is needed to provide a better understanding of the cognitive load associated with the processing of gender and number disagreement.

Often disregarded yet likewise essential for understanding language processing, are the mechanisms of morphosyntactic agreement computation. Such computations are believed to be highly automated, potentially operating in isolation, without reliance on semantic interpretation or processing. Syntax-first models, for instance, advocate for a modular syntactic processing system, which means that in the earliest moments of structure building, only syntactic constraints play a role (Frazier & Coltheart, 1987; Friederici, 2002). Alternative approaches, however, suggest a parallel processing model where lexical and grammatical streams can operate both independently and simultaneously, with the balance between them shifting based on the type of linguistic input and context (Baggio, 2021; Jackendoff, 2007; Kuperberg, 2007). This leads to the following question: would a reader still compute agreement relations if one of the words involved does not actually exist but merely resembles a real word? Pseudowords, which mimic the phonotactic and morphological properties of real words but lack a semantic representation, can serve as valuable tools for such investigations. If syntax is processed autonomously prior to semantics, one might expect agreement computations to proceed unhindered. However, if syntactic operations do not necessarily precede semantic ones at the initial stages of structure building, the presence of pseudowords might disrupt the agreement processing. Therefore, we examine the relationship between semantic and syntactic processing in the brain, potentially offering insights into their dynamic interplay.

1.1 Aims of the study

This thesis explores how the brain processes gender and number (dis)agreement in Norwegian by examining patterns of brain activity. Specifically, we investigate differences in brain responses to gender and number violations in Norwegian predicative adjectival agreement. Additionally, the study examines participants' ability to detect morphosyntactic violations in sentences constructed with pseudowords (i.e. letter sequences that follow rules of phonotactics but have no meaning). This extension can facilitate a more comprehensive understanding of the cognitive mechanisms involved in syntactic agreement processing, including cases when a word lacks meaning.

The study aims to shed light on the neural mechanisms involved in gender and number processing in Norwegian and to provide insights into the processing of grammatical structures in the brain. The study further seeks to explore potential differences in the processing of gender and number categories, for which previous studies yield somewhat inconsistent or contradictory results. Therefore, this study can be informative not only for understanding neurological underpinnings of gender and number agreement processing, but also for deciphering the nature of morphosyntactic agreement processing in a broader sense. It seeks to provide insights into the way grammatical agreement relations are established, offering a clearer understanding of morphosyntactic relations and their role in the meaning composition.

While a wide body of research exists studying agreement processing, for instance, in Spanish, English, or German, no study appears to examine these phenomena in Norwegian. Norwegian, however, is well-suited for the investigation of agreement processing because it encodes both gender and number agreement. Norwegian, like other Scandinavian languages, does not have biological gender, but instead marks grammatical gender, which is not based on the biological sex of the entity being referred to. Regarding number agreement, Norwegian distinguishes between singular and plural forms of nouns. Interestingly, it marks both gender and number predicate agreement, which is not typical for other Germanic languages (Giusti, 2021). This work, thus, focuses on determining whether the two types of abovementioned grammatical relations evoke different responses. By referring to the findings of previous research, Norwegian language findings can be contrasted with the findings for both the Germanic and Romance language families.

The **primary research questions** in this work are:

RQ1: What electrophysiological effects do violations of gender and number agreement evoke in Norwegian native speakers?

RQ2: Are there any differences in the electrophysiological processing of number and gender agreement in Norwegian?

The following **hypotheses** have been formed:

H1: The syntactic processing of gender and number involves distinct processing mechanisms, due to the features' inherent differences at the representational level.

H0: There is no difference in the way gender and number agreement features are processed.

Additionally, **secondary research questions** have been formulated to examine the nature of morphosyntactic agreement in short sentences:

RQ3: Is it possible for a reader to compute agreement relations when a word in a sentence lacks lexical representation?

RQ4: What electrophysiological effects do morphosyntactic violations in short sentences with pseudowords evoke?

This master's thesis aims to investigate the neural correlates of grammatical agreement processing in Norwegian using the event-related potentials (ERP) method. Through the measurement of electrical activity generated by the brain, we can directly monitor language processing in real-time. Focusing on morphosyntactic processing, this work examines sensitivity to gender and number agreement violations in Norwegian native speakers. The study utilizes the violation paradigm, which involves presenting a number of stimuli following a predictable pattern together with instances that occasionally violate this pattern. This is a commonly used paradigm in EEG research which allows to examine differences in brain responses to unexpected or unusual stimuli. These responses are believed to reflect the cognitive processes responsible for detecting and processing different kinds of information, including syntactic and semantic relations.

1.2 Thesis outline

The thesis is outlined as follows. Section 2 presents an overview of the theoretical accounts of number and gender agreement, including a brief discussion of gender and number agreement systems in Norwegian. This section will provide a concise description of the electrophysiological method and a literature review of electrophysiological studies on syntactic and semantic processing. Section 3 provides a detailed description of the current study and its methods, before Section 4 presents the results of the analysis. Lastly, Section 5 contains a discussion of the research findings, limitations of this study, and the conclusion.

2 Background

2.1 Grammatical agreement

In sentences, syntactic rules shape the morphological structure of words by employing inflectional morphology to indicate their grammatical functions (Deutsch & Bentin, 2001). These syntactic constraints, known as agreement rules, encompass different grammatical dependencies, including gender, number, case, person, and definiteness. Although the degree to which languages utilize agreement rules can vary, they are a fundamental principle in the syntax of many languages (Deutsch & Bentin, 2001).

Steel (1978) defines agreement as the “systematic covariance between a semantic or formal property of one element and a formal property of another”. Agreement relations are asymmetrical and thereby presuppose the presence of two elements: the controller and the target. The former refers to the component which determines agreement, while the latter is the component whose form is determined by agreement (Corbett, 2003). For example, in the determiner phrase ‘these houses’, it is the noun ‘houses’ that triggers plural agreement on the demonstrative pronoun ‘these’. It is evident that the morphological property of the pronoun, in this case, is based on the value of the trigger noun. Thereby, determiners and adjectives are considered to be genderless in the lexicon, as they acquire gender only through grammatical agreement or concord (Harris, 1991). Concord is a linguistic phenomenon closely connected to grammatical agreement. Although these terms are related, some scholars argue that grammatical agreement and concord can have fundamentally distinct characteristics (Baker, 2008). In the minimalist framework, the process by which features are transferred from the subject to the inflected verb is commonly referred to as ‘agreement’ (Baker, 2008). In the sentence “The girls are singing”, the plural noun ‘girls’ must agree with the plural verb ‘are’. On the other hand, the transfer of features from a noun to an adjective is typically called ‘concord’. For example, in the phrase ‘nydelige blomster’ (Eng. beautiful flowers), the adjective ‘nydelige’ concords with a plural noun ‘blomster’ in number. Concord specifically involves the alignment of target words to match the properties of their controllers. Therefore, throughout this thesis I adhere to the use of the term "grammatical agreement", also when discussing adjectival agreement. The use of this broader term simplifies the discussion of agreement and ensures coherence within the study. It further allows to maintain a broader psycholinguistic perspective that does not restrict this thesis to a single theoretical framework.

Agreement is a complex phenomenon and its patterns commonly involve the variation of three categories: person, number, and gender (Wechsler, 2009). According to the Feature Hierarchy hypothesis, some categories are more important than others: Person > Number > Gender (Greenberg, 1963). If one of these features is present in a language, then the language also must possess those features hierarchically above it. For instance, if the category of gender is found in a language, then the features of number and person also must be present. Thus, person is the most basic feature that can occur independently of the other two (Molinaro et al., 2011). Furthermore, it has been proposed that the feature hierarchy reflects the degree of the cognitive salience of these features, whereby the features at the top of the hierarchy are more cognitively significant than those at lower levels (Carminati, 2005). In other words, person may impose greater processing costs than number, and number, in turn, may need more processing resources than gender. De Vincenzi (1999), using a priming technique, demonstrates that number and gender information is used at different stages in Italian sentences with object clitic pronouns. Priming effects for the number feature were obtained at the early processing stages, in comparison to semantic gender, which displayed effects only at the end of the sentence. This finding is often used as evidence to show that number processing is costlier than gender processing, as number information is used earlier in processing than gender. However, the results of the study by Silva-Pereyra and Carreiras (2007) do not support the Feature Hierarchy hypothesis. Using event-related brain potentials, Silva-Pereyra and Carreiras (2007) determine whether the morphological feature of person has indeed a different cognitive strength than the number feature. In their experiment, native speakers of Spanish were asked to read sentences in which person agreement, number agreement, or both were manipulated (e.g. Nosotros$_{1stPerPl}$ entiendo$_{1stPerSing}$ la idea [We understand the idea]) (Silva-Pereyra & Carreiras, 2007, p. 207). The Feature Hierarchy hypothesis would predict that disagreement in person would elicit larger effects than number disagreement because the person feature is higher in the hierarchy. However, both conditions showed similar P600 effects. Interestingly, the third condition with combined violation of person and number agreement showed larger amplitudes at the first phase of the P600 than for a violation of person and number separately. The authors suggest that greater effects in the first window of the P600 could be the results of "additive processing effects" (Silva-Pereyra & Carreiras, 2007, p. 206). In other words, the two types of disagreement could require twice the speed and resources as compared to the processing of single features. Concerning the comparison of person and number disagreement, the obtained results suggest no significant difference between the features of person and number in the early phase of processing. As pointed out by Silva-Pereyra and Carreiras (2007), certain differences between person and number may arise at the late stages of processing. Another possibility is that the feature salience can be related to distinct aspects of each language, which would invalidate the hypothesis. Furthermore, other ERP studies contrasting grammatical features do not support the feature hierarchy hypothesis (Barber & Carreiras, 2005; Mancini et al., 2011a; Silva-Pereyra & Carreiras, 2007).

2.2 Categories of gender and number

The complexity of gender lies in the fact that gender assignment is arbitrary, signifying that the meaning of a noun is not always sufficient to determine its grammatical gender. Hence, gender, as an intrinsic feature of nouns, is especially problematic to acquire for non-native speakers. Gender-class information is stored with the stem of the noun in the mental lexicon and must be learned together with the noun (Barber & Carreiras, 2005). However, the acquisition of gender involves not only lexical knowledge but also knowledge of the gender agreement system. Gender assignment may hinge either on the meaning of the noun (semantics) or on its form (Corbett, 1991). Gender can be semantically interpretable when referring to humans or other animal referents with biological sex. However, it becomes entirely grammatical and semantically uninterpretable when referring to inanimate objects (Harris, 1991; Molinaro et al., 2011). Transparency of conceptual gender allows to allocate nouns to genders with ease, while there is no semantic basis for grammatical gender, which is assigned arbitrarily. Furthermore, grammatical gender can be overt or covert. Whereas overt gender can be determined by the form of the noun, covert gender is not evident from the formal features of the noun (Corbett, 1991). In the latter case, gender predictions can be made based on the morphology of the language. Russian is an example of a language with a moderately overt gender system, where morphological features have a crucial role in the gender assignment process (e.g. nouns ending in -a normally belong to the feminine gender).

In contrast to the category of gender, number is a conceptual category, which indicates the plurality of the referent (Corbett, 2000). There is no formal meaning in number, since its interpretation is related to the plurality of the referent, making it more transparent than gender (Arsenijević & Borik, 2020). In other words, number is semantically interpretable on nouns. Number is deemed to be an autonomous feature computed at the syntactic level as opposed to the lexically stored gender feature. Number values of a noun can vary widely across languages. The value of grammatical number usually varies between an unmarked form used for the singular and a marked form encoded for the plural. However, certain languages also implement additional forms to express the presence of three entities or a small number of entities (i.e. paucal forms) (Corbett, 2000). Nonetheless, the category of number is not as transparent as it appears to be. For example, notional number is not always reflected in the grammatical one (Corbett, 2000). In some cases, notionally singular nouns such as ‘scissors, trousers, spectacles’ are grammatically plural. At the same time, some nouns do not mark plural at all. For instance, English abstract nouns such as ‘courage, patience, joy’ have only singular forms.

Thus, both grammatical gender and number appear to be complex categories when it comes to agreement. These features, however, clearly have certain representational differences: whereas number is a computational feature, grammatical gender is a feature of the lexical representation, suggesting the presence of representational differences between the features. Hence, the following question arises: do these representational differences affect the syntactic processing of gender and number agreement?

2.3 Behavioral studies on gender and number agreement

The literature yields contradictory findings regarding differences between gender and number agreement. Language production data, for instance, suggest that gender and number agreement might be processed and represented differently. Specifically, the difference in gender and number agreement error rates could suggest the involvement of different agreement mechanisms. Vigliocco et al. (1996), for example, employ a sentence completion task involving subject-verb agreement in Spanish and English. The authors report that out of 2051 responses, there were only 3 gender agreement errors compared to 103 number agreement errors. This discrepancy might be attributed to the fact that gender is retrieved automatically together with the lemma, in contrast to number, which is processed by means of computation. Consequently, this representational difference indicates that there are fewer chances for an error in gender. Igoa et al. (1999) further investigate speech-error data to determine how gender and number are represented and processed during language production. In their analysis, the researchers discover that exchange errors mostly impact number suffixes, rarely gender and number together, and never gender alone. The results further demonstrate that while grammatical gender is part of the lemma and retrieved directly from the lexicon, number is generated by the rule. Gender morphemes, therefore, should be stranded less frequently than number morphemes, which are expected to be allocated directly to the noun.

Conversely, a study by De Vincenzi (1999) indicates that number agreement processing is cognitively more demanding than gender agreement. As has been previously mentioned, De

Vincenzi (1999) report a significant priming effect for number, which according to the authors, suggests an earlier access for this feature. Meanwhile, a priming effect for gender was found only at the end of the sentence. The authors argue that gender is accessed later because it is linked to lexical processes which take place after phrase-building operations are undertaken.

Faussart et al. (1999), however, suggest that both features are processed quite similarly, with differences becoming apparent only in the later stages of processing. The scholars conducted two experiments to examine the effects of gender and number on lexical decision tasks in Spanish and French. The authors manipulated gender and number agreement in a way that the target word (noun) either agreed with a determiner in both gender and number or disagreed either in gender or number. The researchers found that the reaction times for the incongruent conditions were longer compared to the congruent conditions. Most importantly, the results revealed an interaction between grammatical relation (congruent vs. incongruent) and violation type (number vs. gender), with gender violations being more disruptive than number violations (Faussart et al., 1999). Following earlier ideas by Bradley and Forster (1987), Faussart et al. (1999) propose that a target word is retrieved in a three-stage process, involving: (i) lexical identification, (ii) recognition, and (iii) integration (Fig. 1). The first step includes the process of locating the right lexical entry and its identification, while the relevant lexical content of the entry is accessed in the second stage. The third step involves all post-lexical integration processes related to context, including grammatical agreement. If gender agreement is not detected, the parser would have to go back to the first stage to check whether the right lexical entry had been chosen, as gender is a lexical feature. Number, however, is not a lexical feature and thereby only the syntactic integration processes would be checked in case of inconsistencies in agreement. In other words, lexical information, such as gender, is processed earlier than syntactic information, which includes number. While the parser needs to go back to the initial stage of lexical retrieval to check for gender information, a number mismatch can be resolved at a later stage.

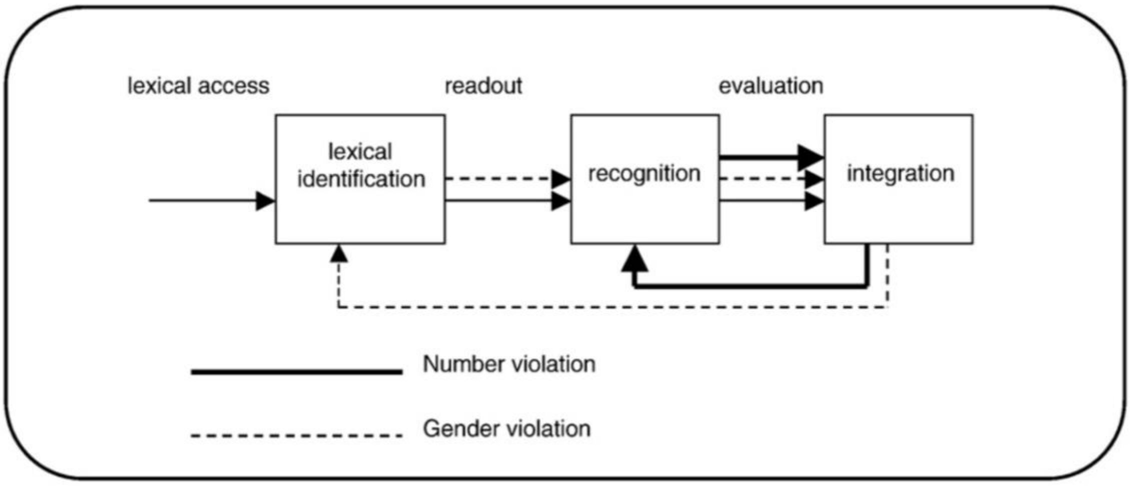

Figure 1: A model of the congruency effect for number and gender proposed by Faussart et al. (1999)

Similarly, Domínguez et al. (1999) explore the way gender and number information can be accessed and processed in the Spanish language. The authors found that when processing words, gender information is primarily retrieved based on the surface frequency of the specific word form, not on the combined frequency of its masculine and feminine forms. In contrast, the processing of plural forms seems more closely related to the frequency of the singular form than to their own frequencies. This suggests a direct retrieval of gender information during word recognition of isolated words, whereas access to number information might involve the lexical entry of the singular form (Domínguez et al., 1999). Acuña-Fariña et al. (2014) utilize an eye-tracking technique to explore the processing of gender and number agreement in the comprehension of subject-verb-adjective sentences in Spanish. One goal of this study was to determine whether there are distinct processing mechanisms for number and agreement errors. In order to answer this question, the authors manipulated attraction/proximity concord in complex subject noun phrases consisting of two nouns (e.g. 'The names of the boy were German' vs '*The name of the boys are German'). The results of this study reveal that participants were sensitive both to number and gender agreement errors. Interestingly, gender agreement anomalies were detected early on at the verb, even though the Spanish verb does not carry gender inflection. While there is a possibility that parafoveal processing is to blame for early gender effects (i.e. the processing of text that occurs just outside the central focus of vision), representation differences could be also accountable for this effect.

Lukatela et al. (1987), nonetheless, found no differences in the way gender and number agreement is processed**.** Particularly, the authors tested Serbo-Croatian native speakers who made lexical decisions on pronoun-noun and pronoun-pseudo-noun pairs. These pairs could agree in case, gender, and number, disagree in either case or gender or number, or disagree simultaneously on two of the three. The findings showed that grammatical congruency influenced both nouns and pseudonouns, with shorter acceptance latencies and longer rejection latencies observed in cases of agreement. Interestingly, the magnitude of the congruency effect was not affected by the type or number of violations. This suggests that the syntactic processing involved, particularly in rapid lexical evaluations, focuses more on the overall grammatical agreement rather than the specifics of the violations (Lukatela et al., 1987).

2.4 Norwegian nominal system

It is critical to understand how grammatical relations work in the language under investigation to accurately manipulate grammatical instances in experiments. Therefore, we discuss the Norwegian nominal system in this subchapter, including details about the constituents involved in the agreement relations, and how such relations operate.

Norwegian, along with Swedish and Danish, belongs to the North Germanic language group. Traditionally, it encodes a three-gender system with masculine, feminine, and neuter genders in most dialects (Rodina & Westergaard, 2015). At the same time, certain dialects distinguish only between two genders: common (fusion of masculine and feminine) and neuter. Norwegian has two written standards: Nynorsk and Bokmål. The standard Nynorsk has all three genders, whereas Bokmål also allows two genders. Moreover, recent studies show that some Norwegian dialects are currently undergoing a gender shift from a three-gender system to a two-gender system (Rodina & Westergaard, 2015). Generally, Norwegian has a relatively complex nominal system, where gender is typically marked on determiners, adjectives, and pronouns, as can be seen below in Examples (2) a-c and (4) a-c.

One significant distinction between Norwegian and other Germanic languages is that Norwegian marks the feature definite on the noun by means of suffixation (Halmøy, 2016). Additionally, Norwegian nouns have a definite plural form which is also marked with a suffix. Thus, most Norwegian nouns have four morphological forms, including Singular Indefinite, Singular Definite, Plural Indefinite and Plural Definite (see Examples (1)).

Examples (1)

a. Et rom – Rommet

‘A room’ – the room’

b. En stol – Stolen

‘A chair – the chair’

c. Spill – Spillene

‘Games – the games’

d. Hunder – Hundene

‘Dogs – the dogs’

Adjectives are inflected to agree with gender, number, and definiteness of the noun they modify. Norwegian provides three forms of adjectives: stor, stort, store. The stem form combines with both masculine and feminine nouns (common gender) and indefinite articles, while the -t suffixed form is used with neuter nouns and indefinite articles (e.g. Examples (2)). The -e form has a broader “distributional and interpretational range” than the other two forms as it is used with all suffixed noun forms: Plural Definite and Indefinite, Singular Definite of all genders (Halmøy, 2016, p. 122) (see Examples (4) a-c and (5) a-c). In contrast to German, where gender and number distinctions disappear with predicative adjectives, Norwegian predicative adjective agreement displays the same inflectional differences as the attributive construction (Vikner, 2009) (e.g. Examples (3) a-c).

Examples (2)

a. en$_{masc,sg,indef}$ stor$_{masc,sg}$ bil$_{masc,sg,indef}$

‘a big car’

b. ei$_{fem,sg,indef}$ stor$_{fem,sg}$ veske$_{fem,sg,indef}$

‘a big bag’

c. et$_{neut,sg,indef}$ stort$_{neut,sg}$ hus$_{neut,sg,indef}$

‘a big house’

Examples (3)

a. Leiligheten$_{masc,sg}$/veska var stor$_{masc/fem,sg}$.

‘The apartment/bag was big.’

b. Huset$_{neut,sg}$ var stort$_{neut,sg}$.

‘The house was big.’

c. Husene$_{neut,pl}$/leilighehene$_{masc,pl}$/veskene$_{fem,pl}$ var store$_{neut/masc/fem,pl}$

‘The houses/apartments/bags were big.’

Examples (4)

a. den$_{masc,sg,def}$ store$_{sg,def}$ bilen$_{masc,sg,def}$

'The big car'

b. den$_{fem,sg,def}$ store$_{sg,def}$ veska$_{fem,sg,def}$

'The big bag'

c. det$_{neut,sg,def}$ store$_{sg,def}$ huset$_{neut,sg,def}$

'The big house'

Westergaard et al. (2017) highlight that gender assignment in Norwegian is usually regarded as being non-transparent, due to the absence of reliable morphological cues for gender. Unlike languages such as Spanish and Russian, where it is often possible to predict gender from morphological endings (e.g. -o for masculine and -a for feminine), Norwegian does not provide similar gender cues. Even though there are certain tendencies that could help to determine the gender of a noun, it must in most cases be learned on a word-by-word basis.

Whereas Norwegian adjectives must agree with nouns in gender, this agreement is limited and can only be applied to adjectives that agree with indefinite neuter nouns (e.g. Examples (2) a-c). When adjectives agree with definite nouns, the e-form is used regardless of gender, meaning that gender in this case is not specified. It should be also noted that there is a considerable fusion of masculine and feminine in adjectives and prenominal determiners. The same forms of adjectives are used with definite/indefinite forms of masculine and feminine nouns (see Examples (2) a-b and (4) a-b).

With respect to number, the Norwegian language distinguishes between singular and plural forms. Plural nouns are generally formed by adding -(e)r to the stem in the indefinite form and -(e)ne for nouns in the definite form (e.g. en bil – biler – bilene). Plurality is marked on adjectives, determiners, and demonstratives. Adjectives are inflected for number by adding the suffix -e to the end of the word, as demonstrated in Examples (5) below:

Examples (5)

a. store$_{pl}$ biler$_{pl,indef}$

'big cars'

b. de$_{pl,def}$ store$_{pl}$ bilene$_{pl,def}$

'the big cars'

c. de$_{pl,def}$ store$_{pl}$ veskene$_{pl,def}$

'the big bags'

The examples above demonstrate that gender is only expressed with singular Norwegian nouns, but not with plural nouns (Jin, 2007). It should be pointed out, that some neuter nouns are never combined with the indefinite plural suffix -er as shown in Examples (6). Consequently, in certain contexts, it may be unclear whether these nouns should be interpreted as singular or plural (Halmøy, 2016). Nevertheless, once a noun is used with a modifier (e.g. an adjective), its interpretation becomes less ambiguous.

Examples (6)

a. Et $\text{hus}_{\text{neut,sg,indef}}$ – $\text{Hus}_{\text{neut,pl,indef}}$

   ‘A house –houses’

b. Et $\text{bord}_{\text{neut,sg,indef}}$ –$\text{Bord}_{\text{neut,pl,indef}}$

   ‘A table– tables’

c. Et $\text{stort}_{\text{neut,sg}}$ $\text{hus}_{\text{neut,sg,indef}}$ – Store_{pl} $\text{hus}_{\text{neut,pl,indef}}$

   ‘A big house – Big houses’

2.5 The ERP paradigm

In this study, event-related potentials (ERPs) are scalp-recorded voltage changes that are time-locked to a specific event or stimuli. These voltage changes are recorded by means of electroencephalography (EEG), which is a non-invasive method used to measure the electrical activity of the brain. ERP recordings are made by putting electrodes on the scalp that detect electrical activity produced by large groups of similarly oriented pyramidal neurons in the cerebral cortex. ERPs are derived from EEG data by averaging EEG samples in multiple trials. The ERP signal is usually described in terms of the amplitude of a specific waveform, the latency of a specific waveform, and scalp topography. This method allows for gathering both quantitative (e.g. positive or negative effects, scalp topography) and qualitative (e.g. amplitude) information about the mechanisms underlying language processing (Luck, 2005). The analysis of ERPs typically involves several steps, including filtering, segmentation of epochs, baseline correction, artifact rejection, averaging, and statistical evaluation (Rommers & Federmeier, 2017).

Language processing is a complex and dynamic process that unfolds in real-time. It involves several cognitive subprocesses, including perception, attention, memory, and decision-making, which work together to enable humans to comprehend and produce language. ERPs have been extensively employed in language research to identify brain activity associated with language processing, in particular with syntactic processing. The multidimensional nature and good temporal resolution of this method make it particularly well suited to the investigation of how agreement relationships are computed during reading. The good temporal resolution allows for recording changes in brain activity in real-time with millisecond precision, which is critical for investigating dynamic brain processes (Luck, 2005). One advantage of ERPs is their ability to show different effects when responding to stimuli that involve grammatical or, for example, semantic anomalies. In contrast, behavioural methods, such as self-paced reading, may not offer the same level of precision in identifying specific cognitive processes. Thus, ERPs have been widely used to study the cognitive and neural processes involved in language processing.

While EEG has several tangible advantages, including its good temporal resolution and non-invasiveness, it also has several limitations, which should be taken into account in experimental design. One disadvantage of this method is its poor spatial resolution, which makes this method unsuitable for investigating processes that require spatial precision (Luck, 2005). EEG signals are created by the simultaneous activity of thousands of neurons, making it hard to pinpoint the source of the activity with great accuracy. Thus, as recommended by Luck (2005), alternative methods such as fMRI should be used when brain topography is considered. Another limitation stems from the fact that EEG is sensitive to different kinds of artifacts, such as eye movements, muscle tension, and electrical noise from external sources. Potential artifacts can obscure brain activity, which makes it difficult to interpret the data. Consequently, in almost all ERP studies subjects are instructed to remain seated in a chair and to maintain constant fixation. While these instructions are necessary to obtain clean and reliable data, they can potentially impact the participant's attention to the stimuli and lead to fatigue. For this reason, the experimental setting is regarded to be highly unnatural, thereby compromising the ecological validity of a study. Additionally, ERPs are relatively small compared to the noise level. Therefore, many trials are normally required to accurately evaluate a specific ERP effect (Luck, 2005). Moreover, as further pointed out by Luck (2005), ERPs as a measure of brain activity are ‘too coarse’ to draw precise conclusions about the underlying brain circuitry. For this reason, it is crucial to make sure that a specific experimental effect has only one possible cause and to ensure that experimental manipulations do not produce unintended side effects that can lead to incorrect conclusions.

Most ERP studies on grammatical agreement processing use the violation paradigm, in which non-grammatical sentences with a specific type of syntactic violation are compared with correct ones, allowing to examine sensitivity to agreement violations. The same paradigm is usually employed for the investigation of semantic processing, where semantic conditions are contrasted to non-semantic ones. The use of the violation paradigm helps to understand how the language processing system recognizes, integrates, or recovers from these violations (Kuperberg, 2007). Two components, the Left Anterior Negativity (LAN) and P600 are typically detected in ERP agreement processing studies. The LAN is deemed to "reflect a stage of processing related to the early detection of a morphosyntactic violation", whereas the P600 component is associated with syntactic violations and thought to reflect reanalysis and syntactic integration processes (Molinaro et al., 2011, p. 915). ERP data indicates that gender and number disagreement produce qualitatively similar effects. The majority of studies on gender agreement processing report a P600 effect, sometimes preceded by a LAN effect, and a similar pattern is reported in most studies on number agreement as well (Barber & Carreiras, 2005; Molinaro et al., 2011; Silva-Pereyra & Carreiras, 2007). The Early Left Anterior Negativity (ELAN) is another component associated with syntactic processing, thought to be involved in the identification of grammatical structures of sentences. The nature and functional role of this component, however, remain a subject of debate. Lastly, the processing of words that are conceptually anomalous often triggers a negative-going wave, known as the N400.

The following discussion illustrates that ERP components demonstrate sensitivity to several factors. It is often challenging to attribute the observed effects to a single factor alone, highlighting the intricate nature of language comprehension and the need for considering the interplay of multiple processes. The sections that follow will provide further detail on each of these potentials.

2.5.1 The ELAN

The Early Left Anterior Negativity (ELAN) is an event-related potential component that occurs early in the timeframe, between 100-300 msec after the stimulus onset. It is predominantly claimed to be an indicator of first-pass syntactic processing, meaning that it is associated with the initial stages of syntactic analysis, i.e. the identification of the grammatical structure of a sentence (Friederici, 2002). The neurocognitive model proposed by Friederici (2002) significantly influenced the interpretation and understanding of the ELAN component within the realm of ELAN research. According to this model, auditory sentence processing consists of three stages manifested in several electrophysiological markers and neurotopographical specifications (see Fig. 2). In the first phase (100-300 msec post-stimulus onset), ELAN is triggered when a word category violation is detected in a sentence. During the next stage (300-500 msec), the parser becomes sensitive to semantic and morphosyntactic violations, manifested in the N400 and the LAN responses, respectively. During the last phase (500-1000 msec), different types of information are integrated, including semantic and syntactic information. If the brain detects sentence structure anomalies, reanalysis and repair processes come into play, reflected in the P600 response. Furthermore, during this stage morphosyntactic violations previously identified in the second phase need to be reanalyzed and repaired, enabling the parser to proceed to the structural integration. Friederici (2002) posits that this model aligns well with both syntax-first models and interactive models that suggest late interaction.

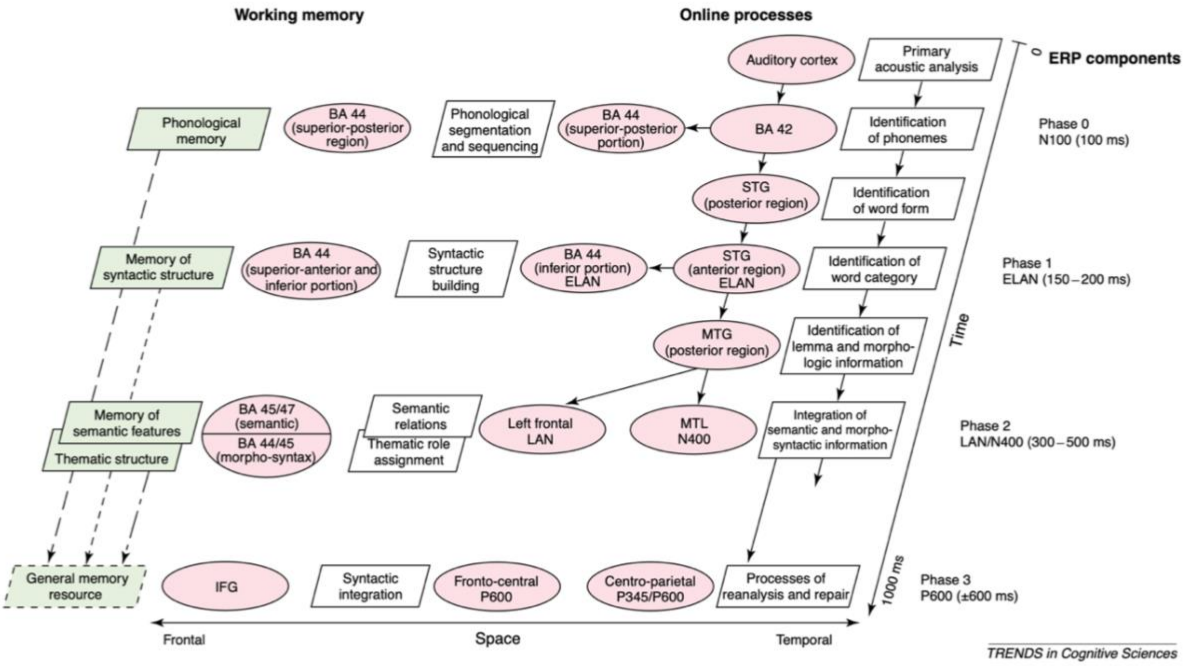

Figure 2: Neurocognitive model of auditory sentence processing proposed by Friederici (2002)

Steinhauer and Drury (2012), however, raise several questions regarding the reliability and validity of the proposed model and the findings of ELAN studies. Firstly, the authors explore differences between audio and visual modalities in the context of ELAN studies. While Friederici's model primarily offers an account of auditory language processing, there is an expectation that ELAN effects should manifest in both domains. ELAN effects, however, are more observed in studies involving connected speech than in studies employing a visual paradigm. The possibility exists that written language, being a learned skill unlike innate speech, might be processed less automatically, possibly impacting the manifestation of ELAN-like effects in reading studies (Steinhauer & Drury, 2012). Thus, the extent to which results from reading studies can be generalized to auditory studies (or vice versa) remains a subject of debate.

Secondly, ELAN studies often use stimuli that do not meet the outright violation criterion. Steinhauer et al. (2008) demonstrate that most sentences utilized by Friederici to establish word category violations can, in fact, be continued to form complete grammatical sentences. Additionally, Steinhauer and Drury (2012) discuss the potential misinterpretation of word category violations as morphosyntactic agreement errors, which typically trigger different ERP components. The mechanism by which readers identify the specific type of anomalies in real-time continues to be unclear. Information about subsequent words, which is crucial for determining the type of anomaly, is not yet available when encountering the target words. The absence of immediate information about subsequent words can lead to crucial interpretations about the type of anomaly encountered, resulting in the potential classification of a sentence as a morphosyntactic violation instead of a word category violation. These observations raise questions regarding the nature of syntactic anomalies and emphasize the importance of accurately categorizing the types of violations presented in ELAN studies.

Furthermore, most phrase structure (PS) violation studies utilize a context manipulation approach, raising additional methodological concerns (e.g. Hahne & Friederici, 1999; Neville et al., 1991). This approach maintains consistency in target words, thereby avoiding potential confounds of violation effects and lexical differences. In such paradigms, words preceding the critical targets systematically vary, leading to potential context effects on target word ERPs, unrelated to the PS violation under investigation. These additional context effects can manifest in auditory studies due to variations in phonological and, notably, prosodic elements, including differences in pitch contour, signal intensity, and duration between violation and control conditions.

Thirdly, Steinhauer and Drury (2012) discuss spill-over effects and offset artifacts, which are context effects that often occur in context-manipulating designs. Spill-over effects arise when the difference in ERP associated with preceding words or stimuli continues to impact the ERP measurements observed for the subsequent target words or stimuli. Essentially, this means that the responses to earlier words have not fully resolved, allowing the effects from the preceding word to extend into the processing of the target word. Two kinds of spill-over effects can be found in ERP studies: (1) effects exhibiting a sustained time course, (2) effects manifesting late in onset. In (1), for example, a late effect such as P600, triggered by words immediately before the target, could potentially be mistaken for an early effect associated with the target word itself. Offset effects, on the other hand, occur when differences in the baseline interval induce erroneous shifts in subsequent ERPs. In such cases, if during the baseline interval the experimental condition is more positive than the control one, the baseline correction will shift the ERP curve, consequently influencing the observed effects. Friederici et al. (1999), for instance, have systematically examined ERP effects for words preceding the critical target words and found marginal ELAN effects, highlighting the susceptibility of ELAN effects to spill-over and offset artifacts. This underlines the importance of detailed analysis of pre-target words to ensure the reliability of the observed effects.

To conclude, given the observed methodological issues in many ELAN studies, careful consideration is necessary when interpreting their results. Numerous studies have significant methodological limitations, pointing to a need for more precise and well-structured research approaches to investigate ELAN effects and to ensure the reliability of the findings. Furthermore, the first phase of (auditory) sentence processing may not be limited only to phrase structure processing. Lastly, the inconsistencies in findings across different ELAN studies indicate that it may be necessary to revise and modify some aspects of Friederici's model.

2.5.2 The LAN

Another event-related potential (ERP) component, the left-anterior negativity (LAN), emerges in an early time window (around 300 ms) and is linked to syntactic processing. The LAN is believed to be associated with morphosyntactic errors, whereas the later P600 correlates with "outright syntactic errors" (Friederici, 2002). The functional nature of the LAN is still debated because the presence and topography of the LAN vary greatly across studies.

A significant number of studies on morphosyntactic processing fails to detect any left anterior negativity (e.g. Deutsch & Bentin, 2001; Hagoort, 2003; Mancini et al., 2011a; Molinaro et al., 2011; Nevins et al., 2007). Moreover, the P600 can sometimes be found in the absence of the earlier LAN, which calls into question the above-mentioned suggestions about the nature of the LAN. Friederici and Weissenborn (2007) propose that the presence of the LAN effects is modulated by the morphological richness of a language: the richer the language is, the more morphosyntactic processing mechanisms need to be employed. However, the LAN effects have also been observed in morphologically poor languages, while absent in some rich ones (e.g. Osterhout & Mobley, 1995, Nevins et al., 2007). Vos et al. (2001) propose that the LAN can be seen as an index of the memory load involved in the detection of an actual violation. The LAN is, thereby, not detected in studies with immediate violations, in which the working memory load is relatively low (e.g. subject-verb agreement). Vos et al. (2001) suggest that the presence or absence of the LAN effect can be influenced by the specific experimental conditions and the level of working memory demand involved. Molinaro et al. (2011) hypothesize that the LAN modulation can be constrained by the salience of cues in the functional morphology of the agreeing constituents. Thereby, the detection of morphologically expressed features may trigger an expectation for the following constituent and its form (e.g. a determiner triggers an expectation for a noun) (Molinaro et al., 2011). Furthermore, the authors point out that the LAN components occur at the same time window as the N400, which is believed to be linked to the lexical predictive processes triggered by semantic information. Nevertheless, driven by a lack of experimental evidence, additional research is needed to establish parallels between these two components. Tanner and Van Hell (2014) further propose that the LAN can be a variant of the N400. The absence of expected robust effects in some studies could be explained by traditional grand averages usually employed to study ERP effects in individuals with otherwise great variability. Therefore, biphasic negative-positive responses may be a by-product of averaging over individuals with different ERP response patterns. While some individuals may focus more on lexical information (N400), others may pay more attention to combinatorial information (P600). The traditional ERP averaging can lead to a misleading interpretation of the results, leading to null-effects or biphasic negative-positive going responses (Tanner & Van Hell, 2014). Tanner and Van Hell (2014) argue that in a normal population the probability of sampling a P600 effect is higher than sampling an N400 effect. Thus, the N400/P600 imbalance is considered to be the cause of the "LAN artifact" in the grand average. On the contrary, Molinaro et al. (2015) argue that if the LAN represents the average response variance to one linguistic manipulation and the N400 represents the average response variance to another, it still supports the conclusion that the population processes these stimuli differently. Moreover, Molinaro et al. (2015) indicate that technical factors, such as the choice of the reference electrode, may influence whether the LAN is present or not. According to the authors, there are higher chances of observing the LAN potential if a hemisphere-independent reference is employed, i.e. the average activity of the two mastoids.

The complex nature of the LAN and mixed research data make it difficult to disentangle what exactly induces the modulation of the LAN effects. For the sake of simplicity, in this work, the LAN is treated as a marker of basic morphosyntactic analysis or identification of morphosyntactic violations.

2.5.3 The N400

The N400 is an ERP component that reaches its peak approximately 400 msec after the onset of a stimulus. In contrast to the P600, it is a negative-going deflection, typically observed at centro-parietal electrode sites. The N400 is commonly linked to the processing of meaning and semantic integration, particularly when a word or sentence violates expectations or does not fit with the context. The N400 is influenced by several factors, including semantic congruity, contextual predictability, semantic priming, and semantic relatedness. It is largely associated with language processing; however, it is not a purely linguistic component. While the N400 is mostly studied in the context of language processing, it has been observed in response to various other types of stimuli, including visual images, gestures, faces, etc. (Schirmer & Kotz, 2003; Sitnikova et al., 2008). Thus, the view of the N400 is shifting from being a language-processing marker to an indicator of meaning processing in a broader sense (Kutas & Federmeier, 2011).

The N400 effect was first discovered in 1980 by Kutas and Hillyard who investigated the influence of sentence context on word recognition. The authors contrasted sentences with semantically congruent (7a) and incongruent endings (7b):

Examples (7)

a. It was his first day at work.

b. He spread the warm bread with socks. (Kutas & Hillyard 1980, p. 102)

Kutas and Hillyard (1980)discovered that semantically incongruent final words elicit a negative wave, which peaks at around 400 msec and has a posterior distribution. Kutas and Hillyard (1980) at first associated this effect with the 'reprocessing' of semantically anomalous information" and labelled it as N400. Thereafter, significant progress has been made in understanding the nature of this measure and the cognitive processes that underlie it. While semantic incongruencies can elicit the N400 component, they are not the only factor for the occurrence of this component. The N400 effect has been also observed for semantically congruent words, with higher amplitudes for words with a low cloze probability (i.e. words that are less predictable from the context) (Holcomb & Neville, 1990). In other words, the more predictable a word is, the smaller the N400 amplitude, meaning that it is easier to integrate the word into the phrase. Besides that, it is also sensitive to word class, exhibiting a greater amplitude for open-class words compared to closed-class words (Kutas & Hillyard, 1983).

Additionally, Van Petten and Kutas (1990) find that the amplitude of the N400 response to content words (such as nouns, verbs, and adjectives) decreases with the linear word position in the sentence, indicating a stronger semantic context established over time. Furthermore, as demonstrated by Federmeier and Kutas (1999), the N400 is influenced not only by the immediate context but also by more general, context-independent information. Specifically, it can be sensitive to the semantic relations between the expected word and the actual word encountered in a sentence. This context can take various forms, including a single word, a sentence, a discourse, or even a non-linguistic context like a sequence of pictures (Kaan, 2007). Thus, the N400 effect has been regarded as "a marker of the difficulty of semantic integration" into the preceding context (Mueller, 2005, p. 156). Therefore, the N400 has been utilized in studies investigating the mechanisms underpinning semantic processing in first and second-language learners (Hagoort et al., 2004; Hahne & Friederici, 2001; Kotz & Elston-Güttler, 2004; Newman et al., 2012).

Due to its semantic nature, the N400 component is normally not elicited by syntactic violations. Nonetheless, Barber and Carreiras (2005) discovered an N400 for gender and number agreement violations in word pairs. While gender and number disagreement in phrases resulted in "an N400-type effect", agreement violations in the same phrases, integrated into sentential context, elicited the predicted LAN-P600 pattern. The authors propose that agreement of word pairs may be processed at the lexical level by assigning word endings to morphemes, rather than at the syntactic level.

2.5.4 The P600

The P600 is a positive going wave that normally occurs at around 500-700 msec over centroparietal regions. The P600 component is typically observed in response to a series of (morpho)syntactic violations, such as phrase structure violations, syntactic ambiguities, violations in the agreement of syntactic features, and thematic roles violations. Therefore, it is thought to reflect cognitive processes of syntactic reanalysis and repair, or, more broadly, to indicate the difficulties of syntactic integration (Friederici, 2002; Kaan et al., 2000; Mueller, 2005). A considerable number of studies report the presence of the P600 component in response to grammatical incongruencies, demonstrating its robustness across various languages. (Deutsch & Bentin, 2001; Kaan et al., 2000; Mancini et al., 2011b; Nevins et al., 2007; Silva-Pereyra & Carreiras, 2007; Vos et al., 2001). Although the P600 is often associated with language processing, P600 effects can be also found in non-linguistic domains, e.g. violations of visual structure narratives, harmonic anomalies in music, etc. (Cohn et al., 2014; Patel et al., 1998). This indicates that the P600 reflects integration difficulties into the structure of the preceding context, regardless of whether this structure is related to syntax or the language (Kaan, 2007; Kaan et al., 2000).

Kutas and Hillyard (1983) were first to discover the P600. In their study, participants were presented with passages containing semantic anomalies and various grammatical errors (e.g. “As a turtle grows its shell grow too”; “Turtles are smarter than most reptiles but not as smart as mammals such as dogs or socks”) (Kutas & Hillyard, 1983, p. 541). The scholars observed an increased N400 for semantic violations and a posterior positivity evoked by syntactic violations. Kutas and Hillyard (1983) conclude that grammatical violations are processed differently compared to semantic deviations, thereby eliciting distinct effects. Osterhout and Holcomb in 1992 first reported the presence of a widely distributed positive wave, which was evoked by syntactically ambiguous sentences (e.g. “The woman struggled to prepare the meal” vs “*The woman persuaded to open the door”). The observed effect had distinct features, manifesting in polarity, onset, duration, and scalp distribution, which differed from the previously observed N400 component. The results of these experiments demonstrate that the P600 effect is different from the responses typically elicited by semantically incongruent words. Osterhout and Holcomb (1992), therefore, suggest that the P600 effect is sensitive to syntactic anomalies, including syntactic garden-path effects. However, in certain cases, agreement violations may not evoke an expected P600. Münte et al. (1997), for instance, find that number agreement violations in sentences with pseudowords failed to evoke a P600 effect. In contrast, the authors discover that sentences with real words evoke a late positivity resembling the P600, while a negativity with an onset latency of 300 msec was observed for the pseudoword condition. Taken together, these findings indicate that the P600 effect will not occur when the full meanings of individual words cannot be accessed. Additionally, Osterhout et al. (1994) demonstrate that the P600 is sensitive to the level of expectancy for syntactic continuations. Specifically, they show that a P600 component is triggered, when readers encounter local syntactical ambiguities. Osterhout et al. (1994) further investigate how readers resolve these ambiguities. Their approach aimed to determine if readers use general rules of sentence structure (a phrase-structure-based minimal attachment principle) or if they rely on specific information about how individual words typically behave in sentences (word-specific subcategorization information). The findings suggest that the “intensity” of the P600 wave is influenced more strongly by word-specific biases than by the general complexity of the sentence. In other terms, the brain seems to rely on its knowledge of specific word behaviours, when resolving ambiguities in sentences, which highlights the nuanced way our brains process language. Interestingly, the results of two experiments by Kaan et al. (2000) indicate that the P600 reflects a process that is not necessarily related to the reanalysis of syntactic errors or syntactically unexpected continuations. The authors observed a P600-effect in sentences with long-distance wh-dependencies, such as "Emily wonders who the performers in the concert imitate for the audience's amusement.". In sentences without such dependencies (e.g. "Emily wonders whether the performers in the concert imitate a popstar for the audience's amusement."), however, a P600-effect was not observed. This effect was also found in sentences that were grammatical and did not contain any garden-path ambiguity, indicating that the P600 is not limited to syntactic reanalysis processes. The authors conclude that the P600 can also reflect difficulties connected to syntactic integration, specifically in the process of establishing syntactic relationships, such as those involving wh-dependencies.

Several studies, additionally, report P600 effects in response to semantic triggers, including semantic associative relationships, animacy and semantic-thematic violations, plausibility, and context (Fritz & Baggio, 2022; Kim & Osterhout, 2005; Kuperberg, 2007; Kuperberg et al., 2003). Although the N400 is typically seen as a marker of lexical or semantic integration, while the P600 is often linked to syntax, there appears to be a certain degree of overlap between the two components, suggesting that both have an intricate connection to language processing.

The intricacy of the P600 component is further driven by the differences in its scalp distribution, which are argued to indicate distinct underlying processes. For instance, Hagoort et al. (1999) suggest that a frontally distributed P600 reflects reanalysis, while a posteriorly distributed P600 indicates repair. Kaan and Swaab (2003), however, argue that the frontal P600 is associated with processing complexity. It is further proposed that the P600 component can be divided into two distinct phases: an early window with a relatively equal anterior-posterior (between 500 and 700 msec) and a late window with a strong posterior distribution (between 700 and 900 msec). In line with this idea, these phases can be sensitive to different (morpho)syntactic factors and have different scalp distributions (Barber & Carreiras, 2005; Hagoort & Brown, 2000; Molinaro et al., 2008). Specifically, the early phase is associated with structural integration processes and has a broad scalp distribution. The second phase, on the other hand, shows a more posteriorly oriented distribution, indicating reanalysis or repair processes (Hagoort & Brown, 2000).

2.6 Sentence wrap-up in ERP research

Sentence wrap-up is a theoretical concept that refers to processes occurring at the end of a sentence or clause. These processes are believed to affect the results when measuring sentence-final words, to the extent that results can become uninterpretable. Many ERP studies avoid this problem by using sentence medial targets, which can be difficult or impossible to implement in some languages. Nevertheless, there is insufficient evidence for such effects elicited by critical words at the end of the sentence.

The primary question at hand is whether the processes carried out at the end of the sentence indeed have a different nature than those occurring within the sentence. Stowe et al. (2018) point out two types of suggestions regarding the sentence-final processes: (1) the existence of a specific stage of processing that entails "wrapping up the sentence", or (2) the presence of certain stages of linguistic processing that can only occur at the end of a clause or sentence. Thereby, certain stages of linguistic processing, such as syntactic processing or the integration of a proposition within a wider context, occur only at the end of a clause or sentence (Stowe et al., 2018). Consequently, Stowe et al. (2018) argue for a stage distinct from other processes occurring throughout a sentence, which may involve a checking step to ensure that all processing is complete. Thereby, this checking procedure would only occur at the end of the sentence and would likely involve neural resources that cannot be utilized before the end of the clause. An alternative perspective on sentence wrap-up is that it encompasses processes that cannot be completed earlier in the sentence, due to processes such as assigning referents to pronouns, forming connections between clauses, or resolving semantic inconsistencies. Behavioral evidence further suggests that something happens at the end of the sentence or clause (Just et al., 1982; Kuperberg et al., 2006; Rayner et al., 2000). Self-paced reading, for instance, usually shows longer reading times at the end of the sentence or clause. The same pattern can be observed when conducting experiments using an eye-tracking technique. This indicates that the clause is treated as a unit that is not broken up during reading. Nevertheless, Stowe et al. (2018) emphasize that similar effects can be seen at clause boundaries, although at lower magnitude. The presence of similar effects at clause boundaries implies that these processes are not specific to the end of the sentence, thereby challenging the above-mentioned views of sentence wrap-up.

Stowe et al. (2018) argue that the so-called "sentence-final wrap-up" dogma has hindered research on important effects at the end of the sentence. In their work, the authors raise important questions regarding the accuracy of the term "sentence wrap-up". Most importantly, with several arguments, they challenge the idea that "sentence wrap-up effects" can negatively impact ERP results. Self-paced reading, eye-tracking, and ERPs have shown that processing at multiple levels occurs instantly, including morphological processes, syntactic integration, resolution of structural ambiguity, and anticipation of upcoming words or structures (Stowe et al., 2018). For example, some evidence does not support the general idea that the establishment of syntactic relations is delayed until the end of the sentence (e.g. Osterhout & Holcomb, 1992; Osterhout & Mobley, 1995). Moreover, the effects that are observed in ill-formed sentences are not always elicited in grammatical sentences, which makes it limited to ungrammatical sentences. The authors also note that the final words in a sentence are processed differently than earlier words when a decision task is used. The effects elicited from decision-making processes, like acceptability judgments, should not be confused with sentence-final processes. Since decision tasks have a different effect on final words than earlier ones, it is necessary to avoid this type of task when critical words are placed at the end of the sentence. Nonetheless, it should be noted that wrap-up effects are typically found in longer linguistic units such as sentences, rather than in short phrases. Short phrases may not provide enough context to elicit wrap-up effects.

In the context of the current study, it should be mentioned that no decision-making tasks are employed in the experiment to avoid any unwanted additional effects. Secondly, wrap-up effects are typically found in more extended textual contexts, rather than in short phrases or sentences, which, again, may not provide enough context to elicit these effects. Readers, thereby, can process short sentences without the typical cognitive load associated with longer, more complex sentences. Lastly, no full stop was used at the end of the sentence, which further minimizes the likelihood of sentence-final effects.

2.7 ERP studies on gender and number processing

Research on language processing has revealed how and which components are activated during different aspects of language comprehension (e.g. Hagoort & Brown, 2000; Kutas & Hillyard, 1980, 1983). While some studies highlight the syntactic nature of number and gender agreement processing (e.g. Osterhout & Mobley, 1995), other studies explore the interplay between syntactic and semantic factors in sentence comprehension, some especially focusing on the qualitative distinctions between gender and number features (e.g. Barber & Carreiras, 2003; Martín-Loeches et al., 2006; Nevins et al., 2007). Despite a large body of research, a complete understanding of the specific processes behind agreement processing has not been achieved, especially because many studies yield inconsistent results.

A study conducted by Osterhout and Mobley (1995), for instance, plays a significant role in confirming the idea that the brain's "response to a variety of syntactic anomalies and agreement violations is dominated by a positive-going wave, whereas a response to semantically anomalous words is dominated by a negative going wave" (Osterhout & Mobley, 1995, p. 760). Moreover, this finding helps to establish that violations in agreement are perceived as syntactic, rather than semantic. In particular, Osterhout and Mobley (1995) examine whether the features of number and gender are linked to syntactic or semantic ERP responses. The authors' first experiment involved the testing of three different conditions in English: (1) subject-verb agreement violations; (2) number agreement violations between a reflexive pronoun and its antecedent; (3) gender agreement mismatches between a reflexive pronoun and its antecedent. As a part of the experiment, subjects were additionally asked to judge the grammaticality of the presented sentences. The findings of this experiment demonstrate that all types of agreement violations triggered a P600 response, indicating difficulties in syntactic processing. The authors interpret this as evidence that number and gender agreement are primarily processed syntactically, rather than semantically. In a second experiment, Osterhout and Mobley (1995) further tested the effects of agreement mismatches involving personal pronouns. To make sure that the P600 component is not associated with an unexpected task-relevant anomaly, the authors included a condition containing sentences with semantically anomalous words. Osterhout and Mobley (1995) find that syntactic violations yield a P600-like positivity, while semantically incongruous words elicit an enhanced N400 response. Lastly, in a third experiment, the authors examined the same effects but without an acceptability judgment task. They incorporated sentences from the two previous experiments, including the subject-verb number sentences, reflexive-antecedent number and gender sentences (experiment 1), and semantically incongruent sentences (experiment 2). This was made to ensure that the elicited effects are not just artifacts of the sentence-acceptability judgment task. In contrast to the first two experiments, reflexive-antecedent number and gender conditions did not reveal any significant effects as compared to the controls, while sentences containing semantically incongruent words evoked an enhanced N400 response. These results indicate that the effects in the reflexive-antecedent conditions might be related to the nature of the behavioural task employed. Specifically, the task could influence the likelihood of anomaly detection in the reflexive-antecedent conditions.

2.7.1 The interplay of syntax and semantics during sentence processing

The studies discussed in this subsection further examine gender and number (dis)agreement, although with a greater focus on the interplay between syntax and semantics. Syntax and semantics are commonly linked to two distinct but interacting levels, each with its unique characteristics and roles. Jackendoff (1999), for instance, proposes a three-level architecture, consisting of separate processing levels for conceptual/semantic information, orthographic/phonological information, and syntactic information. Language comprehension and production involve the coordination and integration of information across these levels. The N400 and P600 effects have been established as indicators of distinct levels of processing, specifically semantic and (morpho)syntactic processing. Based on these findings, researchers seek to understand the interaction between these two levels across different conditions. The question at hand, however, is to which extent and under which conditions these processes interact with each other.

Focusing on this question, Hagoort (2003) explores the interaction between the two semantic and syntactic processing levels during online sentence comprehension. The main objective of this study was to investigate the effects of combined semantic and syntactic violations in relation to single semantic and single syntactic violations in Dutch. In syntactic violations, gender and number agreement was manipulated in noun phrases (article and noun mismatch) either in sentence-internal or sentence-final positions. Semantic violations consisted of semantically incongruous adjective-noun combinations in the same noun phrases. Lastly, combined violations consisted of a combination of the stimuli from both syntactic and semantic conditions. The experimental task involved reading the sentences and determining whether they were acceptable or not. Hagoort (2003) report a classical P600 elicited by grammatical gender and number agreement violations. In this condition, both the acceptability judgments and brain responses demonstrate sensitivity to gender and number mismatches. Additionally, an increased negativity, similar to the N400, was observed for syntactic violations but only in sentence-final positions. In contrast, semantic violations elicited a significant N400 effect, which was more pronounced at the end of sentences than in sentence-internal positions. Hagoort (2003) suggests that this phenomenon can be attributed to the fact that "the strength of semantic constraints increases towards the end of the sentence" (p. 894). Thus, a violation of these constraints likewise has a more pronounced effect at the end of a sentence compared to earlier word positions. With regards to the third condition, combined violations exhibited an increase in the amplitude of the N400 component as compared to purely semantic violations. However, no significant difference in the P600 effect was observed between the combined violation and the single syntactic violation. The combined condition results indicate an asymmetry in the interaction between syntax and semantics. Specifically, while syntactic analysis remains unaffected by semantic integration problems, the process of semantic integration becomes more challenging in the presence of syntactic violations (Hagoort, 2003). Additionally, when examining the semantic and syntactic conditions, the behavioural data revealed that subjects took longer to evaluate semantic acceptability compared to syntactic acceptability. This could be driven by the fact that the criteria for detection of syntactic violations are clear-cut: something is either grammatical or not. In contrast, semantic integration involves a more nuanced and gradual process. In particular, anomaly detection for semantic violations relies "on a measure of processing complexity" and accumulating evidence before determining if something is truly anomalous or not (Hagoort, 2003, p. 896). Lastly, the N400 effects in response to sentence-final words were consistently observed across various types of violations. Regardless of whether the violation was related to syntax, semantics, or other factors, the N400 component exhibited enhanced activity when the violation occurred at the end of the sentence. This pattern can be explained by the presence of so-called wrap-up effects for sentence-final positions. These effects are believed to signify the cognitive processes involved in integrating and finalizing the overall meaning and structure of a sentence (see Section 2.6).

Martín-Loeches et al. (2006) further investigate the electrophysiological processing of number and gender agreement in Spanish. In their experiment, the authors manipulated noun-adjective number or gender disagreement, i.e. morphosyntactic violations, noun–adjective semantic incongruency, i.e. semantic violations, or a combination of both. The primary objective of this study was to examine the interaction between syntax and semantics during sentence processing by examining responses to both semantic and syntactic incongruencies between a noun and an adjective. The results reveal that the impact of a semantic violation on the N400 response remained unchanged even when a syntactic violation was introduced simultaneously. Martín-Loeches et al. (2006) conclude that the process of semantic integration is not influenced by the intactness of syntactic information. Furthermore, the P600 component was elicited by single semantic violations, suggesting that even semantically incongruent information can initiate structural reanalysis. Importantly, the P600 amplitude to combined violations was reduced when compared to purely syntactic violations, demonstrating that semantics and syntax interact at the level of the P600. One explanation for this finding is that less effort may be invested to re-evaluate the syntactic role of the word when processing sentences with semantically unacceptable adjectives (Martín-Loeches et al., 2006). In their conclusion, the authors refer to the work of Kim and Osterhout (2005), who focus on the similarities between language processing and visual processing systems. The visual processing system comprises two parallel streams that exhibit independence in certain aspects, while demonstrating interaction in other aspects. Similarly, in language processing, there are proposals suggesting that syntactic and semantic processing may also operate through independent yet interacting streams (e.g. Baggio, 2021; Jackendoff, 2007; Kuperberg, 2007). These streams exhibit functional independence, due to their ability to recognize attractive analyses specific to each system. However, in the absence of a strongly attractive analysis, the systems can be influenced by other sources of knowledge. Therefore, the functional independence and constant interaction of the syntactic and semantic streams can coexist and are likely to be dependent on the specific circumstances surrounding a particular word in a given sentence (Martín-Loeches et al., 2006).

2.7.2 Comparison of gender and number agreement processing

The following studies focus on comparing the processing of different grammatical agreement features during online language comprehension. These studies aim to investigate how speakers and listeners process and integrate information related to grammatical agreement in real-time using the ERP paradigm. While there is a considerable amount of research on agreement processing, understanding specific processes involved in agreement mechanisms remains incomplete.

By employing the ERP paradigm, Barber and Carreiras (2005) investigate the electrophysiological processing of number and gender agreement in Spanish. The authors seek to empirically examine the representational differences between these two agreement features. To investigate this, the authors compared gender and number violations that occurred in either initial or middle positions. Barber and Carreiras (2005) conducted two experiments, in which gender and number agreement was manipulated either in sentences or short phrases (e.g. "faro–alto", Eng. lighthouse–high; or "El piano estaba viejo y desafinado", Eng. the piano was old and off-key). The results reveal the presence of the P600 and the LAN component in both gender and number conditions. Interestingly, gender agreement violations elicited a larger effect compared to number agreement violations in the second phase of the P600 effect (500-700 msec). This implies that a gender violation might have a higher impact on the later syntactic processes, such as reanalysis. To explain this finding, the authors referred to Faussart's lexical retrieval model (1999) (see Section 2.3). According to this model, when a gender agreement error occurs, the parser needs to revisit the initial stage of lexical retrieval to verify the gender. In contrast, a number disagreement requires the parser to go back to a later stage to check for syntactic information. Consequently, the larger effects observed for gender violations "would reflect the additional cost of going back one more step" (Barber & Carreiras, 2005, p. 138). Interestingly, while both sentential and phrasal contexts exhibited an anterior negativity, isolated word pairs primarily showed an N400 effect Conversely, violations embedded in sentences exhibited a P600 effect, which might be associated with deeper syntactic analysis in sentential contexts. One explanation is that the observed N400 indicates challenges in integrating morphological features, in particular, matching endings of the words.

O'Rourke and Van Petten (2011) further examine two different types of morphosyntactic agreement in Spanish. In their experiments, the authors manipulate distance and agreement type between the agreeing words (e.g. determiners, adjectives and nouns) to see whether distance can influence the way grammatical relations are established. In the first experiment, participants were instructed to look for errors, whereas the second experiment involved only a comprehension task. O'Rourke and Van Petten (2011) report more pronounced LAN and P600 effects for gender violations when compared to number violations. The authors' interpretation is that the number mismatches between noun and determiner or adjective were likely to be entirely overlooked, leading to a reduced amplitude in both components. Additionally, the findings from the error-detection experiment indicate that even when readers were explicitly instructed to actively look for number agreement errors, these violations were less frequently detected in comparison to gender agreement violations. The authors conclude that gender errors appear to be more noticeable than number errors, at least when considering Spanish speakers. However, it should be mentioned that the experiment was not designed to directly compare the two morphological features, with number violations that occurred in earlier portions of their sentences than the gender violations. Barber and Carreiras (2005) note that violations occurring later in sentences elicited larger P600 effects compared to violations occurring earlier, suggesting that the differing saliency of these two types of violations might be attributed to heightened attention to later sentence portions rather than intrinsic distinctions between them.

Several studies, nonetheless, failed to find any differences in the processing of gender and number agreement violations. Nevins et al. (2007) conducted a study that focused on comparing the electrophysiological processing of person, number, and gender in subject-verb agreement in Hindi. Hindi has a rich morphological system where verbs are marked with the person, gender, and number features of the nominative subject noun phrase. Both online and offline tasks were employed to examine differences between the grammatical features. An equally robust P600 response and no LAN were observed for gender and number conditions, signifying that the two features are processed similarly. Violations involving the person feature elicited a larger P600, which was attributed to a greater salience of this feature (Nevins et al., 2007). Moreover, mismatches involving the person feature were identified more accurately and quickly and perceived as less acceptable compared to other agreement violations. Secondly, no additive effects have been found in responses to multiple feature violations. In other words, mismatches including one incorrect feature were processed in the same manner as violations containing two incorrect features. Alemán Bañón (2010), likewise, do not find any differences between gender and number agreement processing. The author conducted a study with Spanish native speakers, focusing on the claims made by Barber and Carreiras (2005) regarding gender and number agreement. The stimuli used in the experiment consisted of sentences containing noun-adjective violations. Contrary to the pattern observed in the work by Barber and Carreiras in 2005, Bañón et al. (2012) find no significant distinctions between the processing of gender and number violations in the later stages of the P600 component, a critical period associated with syntactic reanalysis and repair processes. Both forms of violations evoked similar responses, thus indicating that these grammatical features might be processed in a similar fashion. The author, therefore, suggests this finding is most effectively explained by models that assign equal importance to both number and gender features.

It is worth noting that none of the discussed studies specifically investigated agreement processing in a Nordic language. Furthermore, most of these studies employed a behavioural task, requiring participants to answer additional questions or rate the grammaticality of the sentence. This could potentially influence the way participants reacted and engaged while reading the stimuli, as discussed in Section 2.6. Thus, inconsistencies in findings underscore the need for further research and serve as motivation for this study.

3 Methods

3.1 Participants

Thirty-three native speakers of Norwegian (17 female) took part in the study. All of them speak Norwegian as their first language. Their age ranged from 19 to 44 at the time of the experiment, with an average of 25.03. All participants were right-handed, had normal or corrected eyesight and reported no history of neurological, learning, or psychiatric disorders. Participants were recruited through the distribution of flyers and by word of mouth. Every subject provided written informed consent to participate in the study. Five participants were removed from the final analysis due to a high number of artifacts in the recorded data. All participants received a gift card as compensation for their participation in the study. The study was approved by the Norwegian Center of Research Data (Reference number: 278916).

3.2 Materials

A total number of 480 sentences with the same syntactic structure [N[V Adj]] was created for the experiment. To avoid additional processing costs on working memory, simple syntactic structures were employed. Experimental materials, constructed in Norwegian Bokmål, were reviewed by a native speaker to ensure their accuracy and appropriateness. Half of the nouns used in the stimuli had common gender and the other half was neuter. The same distribution was upheld for filler sentences. 'Sketch Engine' was used to ensure that the frequency of the noun-adjective combinations was taken into account during control measures (Kilgarriff et al., 2014). The materials consist of eight conditions: GWC, GWI, NWC, NWI, GPC, GPI, NPC, NPI, with 60 experimental sentences in each. In the sentences of the GWC, GWI conditions, gender agreement is manipulated by using a predicative adjective that either agrees (GWC) or disagrees (GWI) with the noun in terms of gender. In the NWC and NWI conditions, the predicative adjective either agrees (NWC) or disagrees (NWI) with the noun in terms of number. In the same way, the sentences of the conditions GPC, GPI, NPC, NPI are formed to manipulate gender or number agreement, but with the help of pseudowords. Norwegian pseudonouns were generated following Norwegian rules of phonotactics. In other words, pseudowords contained letter combinations common in Norwegian so that they sounded Norwegian-like. In the pseudoword condition, pseudonouns either matched (50%) or mismatched (50%) adjectives as far as the gender and number were concerned. The use of pseudowords allowed us to observe how Norwegian speakers integrate syntactic information in the absence of meaning since these non-words follow the morphological rules of Norwegian. To create variation in the experimental design, four blocks of sentences were generated by randomizing the order of the same set of sentences within each block. Each subject was presented with one of the four blocks to prevent them from encountering the same sequence of sentences. The experiment was designed with an equal distribution of correct and incorrect sentences to prevent any bias and ensure a balanced assessment. Examples of experimental sentences created for each condition are presented in Table 1.

Condition	**Example**			
GWC	$\text{Huset}_{neut,sg}$	var	$\text{gammelt}_{neut,sg}$.	(Synt +, Sem +)
GWI	*$\text{Bilen}_{masc,sg}$	var	$\text{gammelt}_{neut,sg}$.	(Synt -, Sem +)
GPC	*$\text{Pedlet}_{neut,sg}$	var	$\text{gammelt}_{neut,sg}$.	(Synt +, Sem -)
GPI	**$\text{Pedlen}_{masc,sg}$	var	$\text{gammelt}_{neut,sg}$.	(Synt -, Sem -)
NWC	Husene_{pl}	var	store_{pl}.	(Synt +, Sem +)
NWI	*Huset_{sg}	var	store_{pl}.	(Synt -, Sem +)
NPC	*Nodlene_{pl}	var	store_{pl}.	(Synt +, Sem -)
NPI	**Nodlet_{sg}	var	store_{pl}.	(Synt -, Sem -)

Table 1: Experimental stimuli

3.1.2 Fillers

Additionally, 120 fillers were created and presented together with the stimuli in a randomized fashion. Gender and number agreement in fillers was manipulated with the help of adjectives: correct or incorrect inflection on the adjective. Filler sentences consisted of 4-6 words and made up one-third of the experimental materials. The fillers were constructed so that 50% of the sentences contained a syntactic violation, and 50% of the sentences were grammatically correct. Fillers strategically employed to ensure that the participants were not able to identify the areas of interest. Examples of fillers can be observed in Table 2.

Agreement	**Filler sentences**	
Number	Studenter trenger praktiske$_{pl}$ fag$_{pl}$.	(Synt +)
	*Studenter trenger praktisk$_{sg}$ fag$_{pl}$.	(Synt -)
	Mange leser lokale$_{pl}$ aviser$_{pl}$.	(Synt +)
	*Mange leser lokal$_{sg}$ aviser$_{pl}$.	(Synt -)
Gender	Hun kjøpte en hvit$_{com}$ kjole$_{com}$.	(Synt +)
	*Hun kjøpte en hvitt$_{neut}$ kjole$_{com}$.	(Synt -)
	De fikk et langt$_{neut}$ brev$_{neut}$.	(Synt +)
	*De fikk et lang$_{com}$ brev$_{neut}$.	(Synt -)

Table 2: Filler sentences

3.3 Procedure

The study was conducted in the language laboratory at the Department of Language and Literature at the Norwegian University of Science and Technology. The experiment was run using the software Presentation®, developed by Neurobehavioral Systems. At the beginning of the session, each participant was informed about the procedure and the possibility of withdrawing from the experiment at any time. Participants were asked to sign an informed consent form and fill out a short background questionnaire. This study applied EEG, a non-invasive technique for recording brain electrical activity. During this test, small sensors and gel were applied to the scalp to pick up the electrical signals produced by the brain. During the experiment, the participants were seated comfortably in a chair in a dimly lit, noise-reducing room, approximately 90 cm away from the computer screen. They were asked to read sentences that were displayed word by word in the middle of the screen (word duration: 260 ms). The stimuli consisted of 20 blocks of 18 sentences each. After the completion of each block, participants were given a short break. Additionally, an extended break was offered in the middle of the session to ensure participants retained their concentration. All subjects were instructed not to blink while the phrases were displayed on the monitor, and only to blink when an asterisk appeared. Participants were also asked to avoid contracting the muscles of their face, neck, and tongue. Each session took approximately 90 minutes, including the preparation and experiment. Following the experiments, the participants were briefly interviewed to assess their level of concentration during the experiment.

3.4 EEG data recording and analysis

ERP data was continuously recorded from 32 electrodes placed on the scalp, using the Easycap 32-Channel Standard EEG Recording cap (Fig. 3). The reference electrode was placed on the left mastoid. During the analysis, all EEG channels were re-referenced to the average of the mastoids, additionally using data from the TP10 channel. The EEG data was sampled at 1000 Hz, utilizing a 1000 Hz cut-off filter and a 10s time constant. The impedance was maintained below 1 kOhm at all electrode sites throughout the experiment. The MATLAB toolbox Fieldtrip was used for data analysis (Oostenveld et al., 2011). EEG segments time-locked to the critical word (adjectives) were extracted from the continuous EEG signal, making an epoched EEG set. The EEG epochs were retrieved starting 200 msec before until 800 msec after the onset of the critical word. The 200 msec pre-stimulus period was used as a baseline correction. Artifact rejection was done using two Fieldtrip functions. Firstly, all trials in which the amplitude values exceeded a threshold of +/- 150 microvolts from the baseline were detected and rejected. In the next step, trials containing movements artifacts were identified and removed by analyzing the z-transformed data from the Fp1 and Fp2 channels and subsequently applying a 1-15 Hz band-pass filter during the pre-processing stage. After that, the clean data was filtered with a digital low-pass filter at 30 Hz to remove any remaining muscle artifacts. To obtain ERPs specific to each condition, artifact-free trials were averaged for each participant within each experimental condition.

A nonparametric statistical method developed by Oostenveld et al. (2011) was used for analyzing the EEG data. This method involved grouping adjacent data points exhibiting a significant difference ($p < 0.05$) into clusters. Each identified cluster's p-value was then evaluated using 1000 Monte Carlo simulations, comparing the summed t-values of the cluster to the distribution of summed t-values obtained from the simulations. This method allows for effectively managing the multiple comparisons problem prevalent in ERP data, providing a robust and reliable means to identify real effects in the data. It is important to highlight that we adhered to the standard method for ERP data analysis, employing a widely accepted approach with conventional values for the main parameters without any modifications.

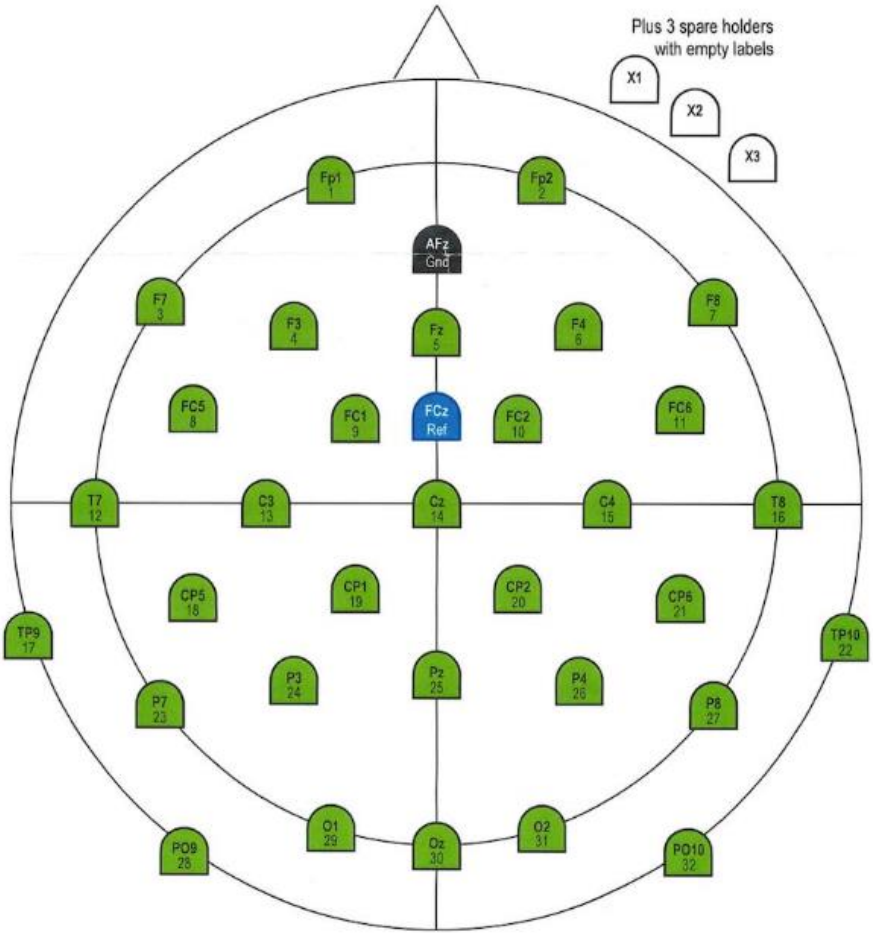

Figure 3: The electrode layout of the Easycap 32-Channel Standard EEG.

3.5 Predictions

Drawing from previous studies on gender and number agreement processing, it is expected that both gender and number agreement violations will elicit a P600 effect, possibly accompanied by a LAN, when compared to their grammatical counterparts (Alemán Bañón, 2010; Bañón et al., 2012; Barber & Carreiras, 2003; Martín-Loeches et al., 2006; Nevins et al., 2007; Osterhout & Mobley, 1995). Regarding differences in gender and number processing, if gender disagreement is indeed more challenging to process, gender violations should evoke a greater positivity than number violations in the second phase of the P600 (Barber & Carreiras, 2005). However, if it is the number that demands more cognitive resources, then number violations should result in a more positive response (Popov, 2022). With regards to the pseudoword conditions, it is expected that violations in non-real words evoke ERP effects, different from those associated with real words. Given that pseudowords are not part of the mental lexicon, their incorrect inflection should not trigger a typical P600 response. It is plausible that the pseudoword condition will trigger an N400 effect, typically associated with semantic anomalies, as it struggles to integrate these non-existent words into a sentence (Kounios & Holcomb, 1994; Münte et al., 1997). Specifically, pseudowords that closely resemble real words may contact semantic memory. However, pseudowords words might not be processed at the syntactic level, due to the lack of lexical representation.

4 Results

4.1 Event-related potentials

The results of the analysis of the ERP data are described below. When comparing gender and number conditions, no significant effects were found. The comparison of the pseudoword condition with the real word condition revealed a larger P600 for the adjective following a real word compared to the adjective following a pseudoword. For incorrectly inflected real words, an early negativity in the range of N100 and P200 time windows was observed. Table 5 provides an overview of the three largest clusters resulting from the nonparametric test.

4.1.1 Effects of features

We discovered no substantial effects attributed to the morphological features in question. Specifically, no significant effects were found in the comparisons between gender incongruent and number incongruent conditions (Fig. 4, row 3), whether involving real words, pseudowords, or both. Similarly, no effects emerged when contrasting number incongruent and number congruent conditions (Fig. 4, row 2) involving any kind of word type. Given these findings, we are unable to reject the null hypothesis, as the statistical analysis revealed no significant effects.

However, in the gender condition, two clusters can be observed (Fig. 4, row 1, and Table 4): 1. A negative cluster (NEG1) within the 100-150 msec range; 2. A negative cluster (NEG2) at around 50 msec. For both clusters, the cluster-level statistics yield a p-value above the 0.05 threshold (NEG1: $p=0.09$; NEG2: $p=0.108$). This indicates that these effects failed to reach statistical significance, thereby suggesting a need for caution in interpretation. The number condition revealed two negativities within the N200 (NEG2) and N400 (NEG1) ranges (Fig. 4, row 2), but this effect was even further from reaching the significance level. Increasing the number of trials or participants might help in determining the true significance of these effects and ensuring their accurate interpretation. There may exist a qualitative difference between the gender and number features, with the effects for the number being notably weaker. Particularly, earlier and larger effects suggest an indication of the incorrectness effect primarily driven by gender, which may be more salient than number. It is important to interpret these findings with caution, emphasizing the necessity for further research to validate these results.

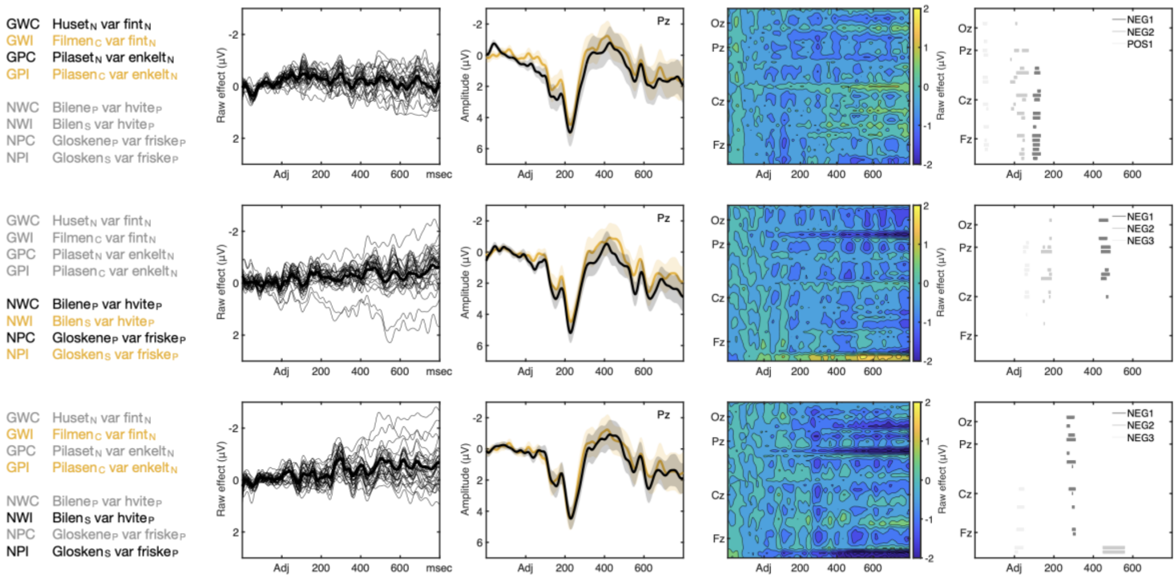

Figure 4: This multi-panel figure illustrates the effects of gender and number features for both real words and pseudowords. Panels in Columns 2 and 5 show the results of the non-parametric cluster-based permutation test. Grand-average ERP waveforms can be observed in Column 3. Column 4 presents the raw effects spanning across various channels and time intervals.

Table 3: Results of cluster-based permutation statistics on ERP data (N=28). The three largest clusters are reported, with the sum of T statistics in each cluster (Tsum), cluster size (S), and the cluster-level Monte Carlo p-value (p).

	Cluster I	**Cluster II**	**Cluster III**
Composition (GWC+NWC)– (GPC+NPC) 0 msec: Adj Figure 5, top	POS1 T_{sum}=3797.3 p=0.025* S=1477	POS2 T_{sum}=705.8 p=0.206 S=2837	NEG1 T_{sum}=-335 p=0.439 S=141
Composition (GWI+NWI)–(GPI+NPI) 0 msec: Adj Figure 5, middle	NEG1 T_{sum}=-456.7 p=0.315 S=180	NEG2 T_{sum}=-214.6 p=0.558 S=91	NEG3 T_{sum}=-152.6 p=0.656 S=66
Composition All W–All P 0 msec: Adj Figure 5, bottom	POS1 T_{sum}=872.8 p=0.149 S=341	POS2 T_{sum}=343.1 p=0.39 S=138	POS3 T_{sum}=-278.8 p=0.451 S=108
Agreement (GWI+NWI)– (GWC+NWC) 0 msec: Adj Figure 6, top	NEG1 T_{sum}=-3829.5 p=0.027* S=1349	NEG2 T_{sum}=-1725.4 p=0.065 S=626	NEG3 T_{sum}=-981.7 p=0.137 S=358
Agreement (GPI+NPI)–(GPC+NPC) 0 msec: Adj Figure 6, middle	POS1 T_{sum}=298.4 p=0.416 S=126	POS2 T_{sum}=177.3 p=0.566 S=76	NEG1 T_{sum}=-118.2 p=0.672 S=53
Agreement All I–All C 0 msec: Adj Figure 6, bottom	NEG1 T_{sum}=-1548.7 p=0.065 S=599	NEG2 T_{sum}=-1506.6 p=0.067 S=569	POS1 T_{sum}=231.1 p=0.464 S=92
Features (GWI+GPI)– (GWC+GPC) 0 msec: Adj Figure 4, top	NEG1 T_{sum}=-1080.2 p=0.09 S=401	NEG2 T_{sum}=-993.3 p=0.108 S=403	POS1 T_{sum}=555.6 p=0.214 S=211
Composition (NWI+NPI)– (NWC+NPC) 0 msec: Adj Figure 5, middle	NEG1 T_{sum}=-625.4 p=0.236 S=267	NEG2 T_{sum}=-463.3 p=0.324 S=190	NEG3 T_{sum}=-322.2 p=0.431 S=140
Composition (GWI+GPI)–(NWI+NPI) 0 msec: Adj Figure 5, bottom	NEG1 T_{sum}=-656.2 p=0.184 S=252	NEG2 T_{sum}=-504.2 p=0.229 S=220	NEG3 T_{sum}=-344.5 p=0.356 S=144

4.1.2 Effects of composition

Using statistical analysis, we examined the differences between the real word and pseudoword conditions, also considering their grammatical correctness. A positive cluster (PO2) was detected in the 100-200 msec time window, although it did not reach the cluster-level significance threshold (*S=2837, p=0.206,* T_{sum}*=705.8*). Secondly, a large positive cluster (PO1) was found between 500 and 700 msec, reaching its peak around 600 msec after the onset of the critical word (*S=1477, p=0.025,* T_{sum}*=3797.32*). This effect has a broad topographical distribution with a larger effect in the posterior region, which is consistent with a P600 effect. The temporal distribution further suggests a P600 effect. Notably, the effect was larger (more positive) for the adjective following a real word than the adjective following a pseudoword. This effect was, however, only found for correctly inflected words as can be seen in Fig. 4. There is no effect for incorrect sentences or both types of sentences collapsed (correct and incorrect), which suggests that the composition process reflected in the P600 is blocked by incorrect morphosyntax.

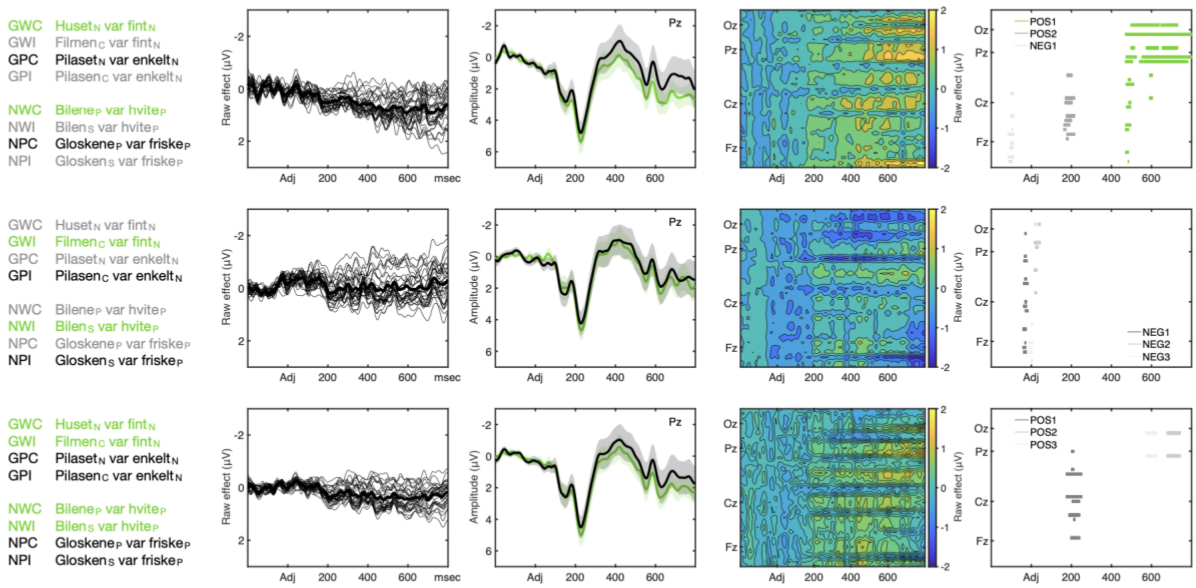

Figure 5: Results of the comparison between real words and pseudowords: non-parametric cluster-based permutation test (Columns 2 and 5); grand-average ERP waveforms (Column 3) and the raw effects over channels and time (Column 4).

4.1.3 Effects of agreement

In the comparison of congruent and incongruent conditions, we observe an early negativity within the N100 and P200 components range. A cluster (NEG1) can be found between 100 and 300 msec post-stimulus onset (*S=1349, p=0.027,* T_{sum}*=-3829.54*). Fig. 5 (row 1, column 3) shows that this negativity peaks around 200 msec after the onset of the adjective. The cluster-level statistical analysis yields a p-value of 0.027. This effect was largest at the occipito-parietal sites (electrodes Oz and Pz), an area which is known as the visual processing area of the brain (Purves et al., 2012). Notably, this negativity was more pronounced for incorrectly inflected adjectives, including both gender and number agreement. This effect is manifested with real words, i.e. when the adjective is preceded and has to agree with a real noun. In contrast, no effect of agreement was found with pseudowords or with real words and pseudowords collapsed. This suggests that agreement processing is triggered only between real words, i.e. items with representations in the lexicon, and it is not a relationship between abstract features independent of lexical items.

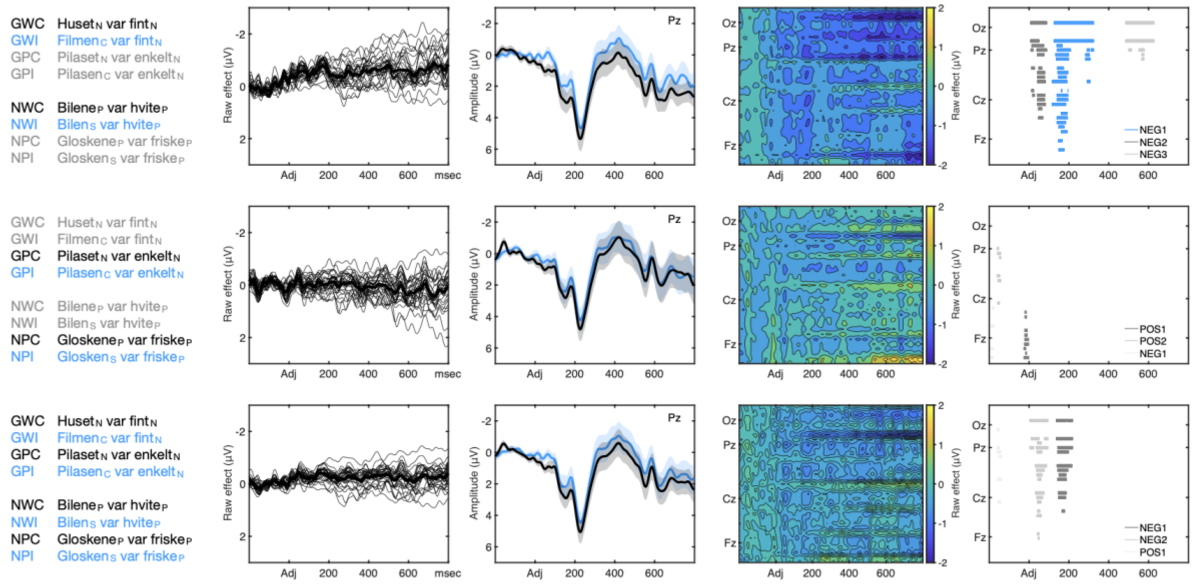

Figure 6: Results of the comparison between congruent and incongruent conditions, involving both real words and pseudoword: non-parametric cluster-based permutation test (Columns 2 and 5); grand-average ERP waveforms (Column 3) and the raw effect over channels and time (Column 4).

5 Discussion

5.1 Gender vs number agreement

The findings related to the gender and number conditions yield results that were, to some extent, unexpected. While previous studies have reported the presence of conventional LAN or/and P600 effects in response to syntactic errors involving gender and number agreement (Barber & Carreiras, 2005; Martín-Loeches et al., 2006; O'Rourke & Van Petten, 2011; Osterhout & Mobley, 1995), our study does not replicate these effects. Instead, we observed an early negativity occurring in the 100-150 msec time window for gender, as well as a hint of negativity in the N400 timeframe for number. These outcomes imply that gender and number agreement violations may be perceived differently by the reader. Nonetheless, this finding should be taken with caution, as the obtained results failed to reach statistical significance, possibly due to a relatively small sample size. However, it is worth noting that the gender effect came close to reaching the threshold of statistical significance and closer than the number effect. This proximity hints at the possibility that gender mistakes may be more salient than number. Several studies have attributed the brain responses elicited within the first 200 milliseconds of word processing to the early word recognition processing (Bentin et al., 1999; Bermúdez-Margaretto et al., 2020; Kong et al., 2012). For instance, research indicates that the N100 component might be linked to the processing of lower-level physical characteristics, such as visual feature analysis (Grossi & Coch, 2005). Therefore, one possible explanation for our results could be that the mismatch in gender might be visually more salient than the mismatch in number, especially given that gender effects manifest in an early time range.

Given that our study does not reject the null hypothesis, the obtained findings do not align with either Fausart's processing model or Popov's assertion regarding the salience of the number feature. The absence of effects similar to those from prior studies may be in part attributed to the difference in the stimuli used in the experiments. Our research utilized shorter sentences that could be categorized as minimal phrases, leaving participants with minimal contextual cues when processing the sentences. The limited syntactic structure of the stimuli might have affected the way agreement relations were perceived and processed. It is also possible that participants engaged in a more ‘superficial’ processing mode, due to the nature of the task. Specifically, their attention might have shifted primarily to the inflection of the adjectives.

5.2 Real words vs pseudowords

The results obtained in the comparison between the conditions with real and pseudowords exhibit similar composition effects as in the two studies by Fritz and Baggio (2020, 2022). When comparing semantic and non-semantic conditions, both early and late positivities were observed. The early P200 effect, however, failed to reach the significance threshold. Walla et al. (1999) posit that the early potential P200 might be linked to the attentional mechanisms underpinning word encoding and recognition. Fritz and Baggio (2022) also suggest that the P200 effect may be indicative of attention modulations between semantic and non-semantic trials. In particular, participants appear to reallocate their attention differently when processing the noun, based on whether it is preceded by a real adjective, by a nonword, or by pseudoword, resulting in variations in attention levels – ranging from normal to heightened or decreased. An alternative explanation is that the early positivity might be linked to orthographic processing, i.e. processes related to the extraction of visual features and word form (Stuellein et al., 2016). Dambacher et al. (2006) find word frequency effects in the P200, proposing that orthographic representations are accessed in this time window.
Therefore, the observed early positivity in our study could reflect the ease or depth with which participants access the orthographic representations of words. The participants' ability to access these representations might in turn be based on their familiarity with the word or its semantic content.

This study replicated the findings of Fritz and Baggio (2020, 2022) by likewise eliciting the P600 effect for composition. Similarly, the effect was more pronounced when the adjective followed a real word compared to sentences in which the adjective followed a pseudoword. Fritz and Baggio (2020) suggest two possible explanations: the P600 can reflect either semantic or both semantic and syntactic composition. Although the pseudowords used in the experiment resemble real Norwegian words both phototactically and morphologically, they are not real lexical units. The resemblance to real words might provide syntactic cues allowing for syntactic composition to take place. However, the absence of a concrete semantic representation could hinder semantic composition. If syntactic computation still takes place, then the P600 can reflect semantic composition with larger effects for real words, due to the presence of actual meaning. In the second scenario, the presence of pseudowords could potentially disrupt both syntactic and semantic composition, meaning that the P600 could reflect both types of processing (Fritz & Baggio, 2020). Kuperberg (2007) hypothesizes that comprehension is driven by multiple distinct but interconnected processing streams. On one hand, there is a system that connects incoming semantic data with pre-existing knowledge stored in the semantic memory. On the other, there is a combinatorial process that is associated with morphosyntactic and certain semantic-thematic constraints. According to this perspective, the P600 can reflect a conflict arising from the conflicting output between these processing streams. Our findings further corroborate the idea that the observed P600 component reflects compositional processes at the syntax-semantics interface.

Concerning the absence of effects with incorrectly inflected adjectives, the results imply that the composition process is blocked due to the presence of morphosyntactic errors. Interestingly, some kind of compositional processes appear to occur even when words do not have any real meaning, while no composition takes place if certain morphosyntactic constraints are violated. This suggests that the compositional processes reflected in the P600 effect might be particularly sensitive to morphosyntactic incongruencies, at least within the context of minimal phrases.

5.3 Congruent vs incongruent

The results obtained for the comparison of congruent and incongruent conditions indicate that the brain fails to establish agreement relations in the presence of pseudowords. While pseudowords are constructed to resemble real words, they lack lexical representations. Consequently, the brain struggles to extract the orthographic representation of the word, i.e. to recognize the word form, which possibly hinders further processing of the unit. In other words, readers try to recognize a word, which does not exist, in their mental lexicon. This interrupts the usual flow of processing, leading to a breakdown in establishing syntactic and semantic relationships within the given context.

The early effects observed with real words also suggest that participants' responses might reflect orthographic word processing, with incorrect forms being more disruptive for a reader. The negativity was larger for incorrectly inflected adjectives, including both gender and number agreement. This suggests that the error detection, in this case, might be manifested in an enhanced negativity. Furthermore, it should be noted that several studies on syntactic processing report similar effects (N100 and/or P200) in response to syntactic violations (Hagoort & Brown, 1999; Hasting & Kotz, 2008; Yamada & Neville, 2007). However, the authors, to the most part, do not provide a comprehensive explanation for these effects.

Since the N100 and P200 are primarily associated with the processing of sensory aspects of experience, it is possible that the effects found in our study are associated with the visual processing of stimuli rather than syntactic processing. In the context of language comprehension, processing stages earlier than 200 msec are thought to be associated with pre-lexical processes, and are thereby not sensitive to 'deeper' lexical or semantic features (Bentin et al., 1999). The magnitude of the P200 component is believed to indicate the differences in accessing the orthographic structure during the recognition of visual words (Bermúdez-Margaretto et al., 2020). This means that the P200 is likely to reflect the initial stages of word form recognition/lexical perception, being sensitive to the visual attributes of the word. This observation stands in line with the findings of Meng et al. (2008), who show that character violations elicit an N100-P200-N400 pattern. Brandeis and Lehmann (1994) further link N100-P200 effects to perceptual processing.

It appears that the early processing phase is important for the subsequent stages of word comprehension, where deeper lexical and semantic analyses take place. If a mismatch or inconsistency is detected at this early stage, as in the case with pseudowords, it can disrupt the typical processing progression. The spatial distribution further supports a linking of the observed effects to visual/perceptual word processing, which aligns with the visual modality of the experiment. Specifically, these effects are most pronounced at the occipito-parietal sites. The occipital region plays a role in primary integration processes, such as the visual perception of forms, shapes, and colors, whereas the parietal area is associated with higher-order processes, like object and form recognition (Purves et al., 2012).

Lastly, the fact that agreement errors are not perceived as grammatical violations might be due to the minimal contextual cues in the stimuli, especially given the simple and consistent syntactic structure of our stimuli. Instead, they might be interpreted as deviations from the expected word forms. In the absence of contextual information, it appears that readers exhibit reduced sensitivity to syntactic anomalies, as this sensitivity is typically reflected in the P600 component.

5.4 Limitations

This study and its approach are subject to several potential limitations. Most limitations, however, originate from the utilization of the EEG method and are unavoidable, due to inherent methodological characteristics.

First and foremost, the experiment was conducted in a very controlled environment, which could induce unnatural behaviour patterns among the participants. In particular, participants were sitting in a dark noise-reducing room, required to avoid any kind of movements throughout the trial to ensure the accuracy and reliability of the EEG data. In addition to this, the stimuli were presented without any context and had a largely identical structure. Such unnatural conditions could potentially make the results less generalizable, as the controlled environment might not reflect real-life situations. Secondly, the study did not gather any additional information regarding individual differences between participants. Including data on factors such as attention span, memory, motivation, or language skills could have refined the results further. Another limitation is the limited scope of the study, as it focuses on a specific instance of grammatical agreement, rather than adopting a more holistic approach. As mentioned earlier, more participants need to be tested to enhance the reliability of the findings, possibly enhancing the significance of the findings. Additionally, most of the participants were students, which may restrict the applicability of the findings to the broader population. Some of these limitations, however, are inevitable in a master's thesis, given the restricted timeframe and resources.

6 Conclusion

In this study, we explored the neural correlates of gender and number agreement processing. By utilizing adjectives that either agreed or disagreed with the noun in number or gender, we aimed to contrast how these grammatical mismatches are processed by Norwegian speakers. Our experimental objectives were twofold. Firstly, we aimed to investigate how grammatical gender and number representations impact syntactic processing, specifically whether violations in number or gender agreement would induce higher processing costs. Secondly, we used pseudowords to test whether these agreement relations can be established in the absence of meaning. Moreover, our experiment did not implement any additional behavioral task. This allowed us to avoid unwanted extraneous effects, thereby giving a clearer perspective on the neurophysiological processes involved in agreement processing.

Regarding the gender condition, we identified a hint of an early negativity within the 100-150 msec time window. The effect found for the number condition was relatively subtle and was observable within the N400 time window. Importantly, neither condition exhibited the presence of LAN or ELAN components. We hypothesize that the processing of grammatical violations in these two features involves different processing routines, due to the distinct representational properties of each feature. However, given the present data, it remains challenging to draw definitive conclusions related to repair or reanalysis processes. Although both effects failed to reach the significance threshold, the results could possibly indicate that these features differ in their visual salience. Specifically, the negativity, observed for the gender condition in the early window, suggests that gender violations might be more prominent at the orthographic or visual processing level than number violations. However, more research is needed to explore the representational differences between these two features and the involved processing mechanisms.

Although the results obtained for gender and number conditions do not let us draw firm conclusions about the way these features are processed, comparison across various conditions yielded interesting findings. For example, we found that agreement relations can be established effectively only between real words. As pointed out by Kim and Lai (2012), visual word recognition requires multiple levels of analysis and representation, including word form processing, phonological, semantic, and syntactic representations. Possibly, when encountering pseudowords, the orthographic representation cannot be accessed, disrupting further processing. Thereby, agreement relations cannot be computed based solely on the syntactic forms of the words, which raises questions about the validity of purely syntax-driven models of language processing. The data also suggest that for real words, processing of incorrectly inflected adjectives may be more challenging already at the visual level, given the more pronounced effects observed for incongruent sentences. The spatial and temporal patterns observed within the N100 and P200 range further support the view that the detected effects are associated with the visual processing of words.

Moreover, the comparison of the conditions with real words and pseudowords revealed a P600 effect, which is believed to reflect composition in minimal phrases (Fritz & Baggio, 2020, 2022). A larger effect was found for sentences in which an adjective followed a real word, compared to sentences in which it follows a pseudoword. One explanation is that compositional processes might be impeded by a conflict between semantic and syntactic information, resulting in less pronounced effects. Specifically, the conflict between syntactic and semantic information arises due to the absence of meaning in pseudowords. The fact that the P600 is observed when comparing semantic (real words) and non-semantic conditions (pseudowords), further indicates that this component does not only reflect syntactic operations (i.e. integration or reanalysis), but rather combinatory analysis at the syntax-semantic interface. It should be also noted the observed effects are not induced by task-related cognitive load, since the experiment did not involve any kind of an additional task (e.g. grammaticality judgement task). Lastly, the absence of effects in the incongruent condition indicates that morphosyntactic violations can block the composition of a phrase. This suggests that in the context of minimal phrases, semantic operations might not take place unless certain morphosyntactic constraints are satisfied. In this way, compositional processes appear to be sensitive to morphosyntactic information (violations), which further supports the idea that the P600 might reflect syntax-driven meaning composition.

As discussed in this study, the minimum context design could have impacted how the agreement relations were processed. It is therefore possible that testing similar conditions embedded in a more extended and natural context can provide more insights into how gender and number agreement is processed. Secondly, increasing the number of participants or trials could further enhance the reliability of the results, especially those for the gender and number conditions. In addition, a multifaceted approach, combining various methodologies and techniques should be employed to understand processes related to morphosyntactic agreement and composition in minimal phrases. In future studies, emphasis should be further placed on the early components that are elicited in sentences with syntactic and semantic violations (e.g.

N100; P200). Specifically, it is important to understand which factors trigger these components in the context of syntactic and semantic incongruences. This could involve factors, such as the type of violation, the complexity of the sentence structure, the cognitive demands, or the nature of the task at hand. Lastly, the idea that the P600 component reflects compositional processes at the syntax-semantics interface should be further tested in various conditions.

References

Acuña-Fariña, J. C., Meseguer, E., & Carreiras, M. (2014). Gender and number agreement in comprehension in Spanish. *Lingua*, *143*, 108-128.

Alemán Bañón, J. (2010). *The processing of number and gender agreement in Spanish: An ERP investigation* University of Kansas].

Arsenijević, B., & Borik, O. (2020). Gender and number in the nominal domain: An introduction. *Catalan journal of linguistics*, *19*, 0005-0024.

Baggio, G. (2021). Compositionality in a parallel architecture for language processing. *Cognitive Science*, *45*(5), e12949.

Baker, M. C. (2008). *The syntax of agreement and concord* (Vol. 115). Cambridge University Press.

Bañón, J. A., Fiorentino, R., & Gabriele, A. (2012). The processing of number and gender agreement in Spanish: An event-related potential investigation of the effects of structural distance. *Brain research*, *1456*, 49-63.

Barber, H., & Carreiras, M. (2003). Integrating gender and number information in Spanish word pairs: An ERP study. *Cortex*, *39*(3), 465-482.

Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of cognitive neuroscience*, *17*(1), 137-153.

Bentin, S., Mouchetant-Rostaing, Y., Giard, M.-H., Echallier, J.-F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: time course and scalp distribution. *Journal of cognitive neuroscience*, *11*(3), 235-260.

Bermúdez-Margaretto, B., Beltrán, D., Shtyrov, Y., Dominguez, A., & Cuetos, F. (2020). Neurophysiological correlates of top-down phonological and semantic influence during the orthographic processing of novel visual word-forms. *Brain sciences*, *10*(10), 717.

Bradley, D. C., & Forster, K. I. (1987). A reader's view of listening. *Cognition*, *25*(1-2), 103-134.

Brandeis, D., & Lehmann, D. (1994). ERP mapping: A tool for assessing language disorders? In *Cognitive electrophysiology* (pp. 239-247). Springer.

Carminati, M. N. (2005). Processing reflexes of the Feature Hierarchy (Person> Number> Gender) and implications for linguistic theory. *Lingua*, *115*(3), 259-285.

Cohn, N., Jackendoff, R., Holcomb, P. J., & Kuperberg, G. R. (2014). The grammar of visual narrative: Neural evidence for constituent structure in sequential image comprehension. *Neuropsychologia*, *64*, 63-70.

Corbett, G. G. (1991). *Gender*. Cambridge University Press.

Corbett, G. G. (2000). *Number*. Cambridge University Press.

Corbett, G. G. (2003). Agreement: terms and boundaries. The Role of Agreement in Natural Language. Proceedings of the 2001 Texas Linguistic Society Conference, Austin, Texas.,

Dambacher, M., Kliegl, R., Hofmann, M., & Jacobs, A. M. (2006). Frequency and predictability effects on event-related potentials during reading. *Brain research*, *1084*(1), 89-103.

De Vincenzi, M. (1999). Differences between the morphology of gender and number: Evidence from establishing coreferences. *Journal of Psycholinguistic Research*, *28*(5), 537-553.

Deutsch, A., & Bentin, S. (2001). Syntactic and semantic factors in processing gender agreement in Hebrew: Evidence from ERPs and eye movements. *Journal of memory and language*, *45*(2), 200-224.
Domínguez, A., Cuetos, F., & Segui, J. (1999). The processing of grammatical gender and number in Spanish. *Journal of Psycholinguistic Research*, *28*, 485-498.
Dowens, M. G., Vergara, M., Barber, H. A., & Carreiras, M. (2010). Morphosyntactic processing in late second-language learners. *Journal of cognitive neuroscience*, *22*(8), 1870-1887.
Faussart, C., Jakubowicz, C., & Costes, M. (1999). Gender and number processing in spoken French and Spanish. *Rivista di Linguistica*, *11*(1), 75-101.
Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of memory and language*, *41*(4), 469-495.
Frazier, L., & Coltheart, M. (1987). *Sentence processing: A tutorial review*.
Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in cognitive sciences*, *6*(2), 78-84.
Friederici, A. D., & Weissenborn, J. (2007). Mapping sentence form onto meaning: The syntax–semantic interface. *Brain research*, *1146*, 50-58.
Fritz, I., & Baggio, G. (2020). Meaning composition in minimal phrasal contexts: Distinct ERP effects of intensionality and denotation. *Language, Cognition and Neuroscience*, *35*(10), 1295-1313.
Fritz, I., & Baggio, G. (2022). Neural and behavioural effects of typicality, denotation and composition in an adjective–noun combination task. *Language, Cognition and Neuroscience*, *37*(5), 537-559.
Giusti, G. (2021). Adjectival Concord in Romance and Germanic.
Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. *Universals of language*, *2*, 73-113.
Grossi, G., & Coch, D. (2005). Automatic word form processing in masked priming: an ERP study. *Psychophysiology*, *42*(3), 343-355.
Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of cognitive neuroscience*, *15*(6), 883-899.
Hagoort, P., & Brown, C. M. (1999). Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of Psycholinguistic Research*, *28*, 715-728.
Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia*, *38*(11), 1531-1549.
Hagoort, P., Brown, C. M., & Osterhout, L. (1999). The neurocognition of syntactic processing. *The neurocognition of language*, 273-316.
Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *science*, *304*(5669), 438-441.
Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of cognitive neuroscience*, *11*(2), 194-205.
Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, *4*(2), 123-141.
Halmøy, M. (2016). *The Norwegian nominal system: a neo-saussurean perspective* (Vol. 294). Walter de Gruyter GmbH & Co KG.
Harris, J. W. (1991). The exponence of gender in Spanish. *Linguistic inquiry*, *22*(1), 27-62.

Hasting, A. S., & Kotz, S. A. (2008). Speeding up syntax: On the relative timing and automaticity of local phrase structure and morphosyntactic processing as reflected in event-related brain potentials. *Journal of cognitive neuroscience*, *20*(7), 1207-1219.
Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and cognitive processes*, *5*(4), 281-312.
Igoa, J. M., García-Albea, J. E., & Sánchez-Casas, R. (1999). Gender-number dissociations in sentence production in Spanish. *Rivista di Linguistica*, *11*(1), 163-196.
Jackendoff, R. (1999). The representational structures of the language faculty and their interactions. *The neurocognition of language*, 37-79.
Jackendoff, R. (2007). A parallel architecture perspective on language processing. *Brain research*, *1146*, 2-22.
Jin, F. (2007). *Second language acquisition and processing of Norwegian DP internal agreement*. Det historisk-filosofiske fakultet.
Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of experimental psychology: General*, *111*(2), 228.
Kaan, E. (2007). Event-related potentials and language processing: A brief overview. *Language and linguistics compass*, *1*(6), 571-591.
Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and cognitive processes*, *15*(2), 159-201.
Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of cognitive neuroscience*, *15*(1), 98-110.
Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography*, *1*(1), 7-36.
Kim, A., & Lai, V. (2012). Rapid interactions between lexical semantic and word form analysis during word recognition in context: Evidence from ERPs. *Journal of cognitive neuroscience*, *24*(5), 1104-1112.
Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of memory and language*, *52*(2), 205-225.
Kong, L., Zhang, B., Zhang, J. X., & Kang, C. (2012). P200 can be modulated by orthography alone in reading Chinese words. *Neuroscience letters*, *529*(2), 161-165.
Kotz, S. A., & Elston-Güttler, K. (2004). The role of proficiency on processing categorical and associative information in the L2 as revealed by reaction times and event-related brain potentials. *Journal of Neurolinguistics*, *17*(2-3), 215-235.
Kounios, J., & Holcomb, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 804.
Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain research*, *1146*, 23-49.
Kuperberg, G. R., Kreher, D. A., Goff, D., McGuire, P. K., & David, A. S. (2006). Building up linguistic context in schizophrenia: evidence from self-paced reading. *Neuropsychology*, *20*(4), 442.
Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, *17*(1), 117-129.
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.

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

Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & cognition, 11*(5), 539-550.

Luck, S. J. (2005). An introduction to event related potentials and their neural origins. *An introduction to the event related potential technique*, *11*.

Lukatela, G., Kostić, A., Todorović, D., Carello, C., & Turvey, M. T. (1987). Type and number of violations and the grammatical congruency effect in lexical decision. *Psychological Research*, *49*(1), 37-43.

Mancini, S., Molinaro, N., Rizzi, L., & Carreiras, M. (2011a). A person is not a number: Discourse involvement in subject–verb agreement computation. *Brain research, 1410*, 64-76.

Mancini, S., Molinaro, N., Rizzi, L., & Carreiras, M. (2011b). When persons disagree: an ERP study of Unagreement in Spanish. *Psychophysiology*, *48*(10), 1361-1371.

Martín-Loeches, M., Nigbur, R., Casado, P., Hohlfeld, A., & Sommer, W. (2006). Semantics prevalence over syntax during sentence processing: A brain potential study of noun–adjective agreement in Spanish. *Brain Research*, *1093*(1), 178-189.

Meng, X., Jian, J., Shu, H., Tian, X., & Zhou, X. (2008). ERP correlates of the development of orthographical and phonological processing during Chinese sentence reading. *Brain research*, *1219*, 91-102.

Molinaro, N., Barber, H. A., Caffarra, S., & Carreiras, M. (2015). On the left anterior negativity (LAN): The case of morphosyntactic agreement. *Cortex*, *66*(156-159).

Molinaro, N., Barber, H. A., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex*, *47*(8), 908-930.

Molinaro, N., Vespignani, F., & Job, R. (2008). A deeper reanalysis of a superficial feature: An ERP study on agreement violations. *Brain research*, *1228*, 161-176.

Mueller, J. L. (2005). Electrophysiological correlates of second language processing. *Second Language Research*, *21*(2), 152-174.

Münte, T. F., Matzke, M., & Johannes, S. (1997). Brain activity associated with syntactic incongruencies in words and pseudo-words. *Journal of cognitive neuroscience*, *9*(3), 318-329.

Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of cognitive neuroscience*, *3*(2), 151-165.

Nevins, A., Dillon, B., Malhotra, S., & Phillips, C. (2007). The role of feature-number and feature-type in processing Hindi verb agreement violations. *Brain research*, *1164*, 81-94.

Newman, A. J., Tremblay, A., Nichols, E. S., Neville, H. J., & Ullman, M. T. (2012). The influence of language proficiency on lexical semantic processing in native and late learners of English. *Journal of cognitive neuroscience*, *24*(5), 1205-1223.

O'Rourke, P. L., & Van Petten, C. (2011). Morphological agreement at a distance: Dissociation between early and late components of the event-related brain potential. *Brain research*, *1392*, 62-79.

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.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of memory and language*, *31*(6), 785-806.

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. *Journal of experimental psychology: Learning, memory, and cognition*, *20*(4), 786.
Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of memory and language*, *34*(6), 739-773.
Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of cognitive neuroscience*, *10*(6), 717-733.
Pesetsky, D., & Torrego, E. (2007). The syntax of valuation and the interpretability of features. *Phrasal and clausal architecture*, 262-294.
Popov, S. (2022). *Auditory and visual ERP correlates of gender agreement processing in Dutch and Italian* Macquarie University].
Purves, D. E., Augustine, G. J., Fitzpatrick, D. E., Hall, W. C., LaManita, A.-S. E., & White, L. E. (2012). Neuroscience.
Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology: Section A*, *53*(4), 1061-1080.
Rodina, Y., & Westergaard, M. (2015). Grammatical gender in Norwegian: Language acquisition and language change. *Journal of Germanic Linguistics*, *27*(2), 145-187.
Rommers, J., & Federmeier, K. D. (2017). 13 Electrophysiological Methods. *Research methods in psycholinguistics and the neurobiology of language: A practical guide*, 247.
Schirmer, A., & Kotz, S. A. (2003). ERP evidence for a sex-specific Stroop effect in emotional speech. *Journal of cognitive neuroscience*, *15*(8), 1135-1148.
Silva-Pereyra, J. F., & Carreiras, M. (2007). An ERP study of agreement features in Spanish. *Brain research*, *1185*, 201-211.
Sitnikova, T., Holcomb, P. J., Kiyonaga, K. A., & Kuperberg, G. R. (2008). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of cognitive neuroscience*, *20*(11), 2037-2057.
Steinhauer, K., Connolly, J. F., Stemmer, B., & Whitaker, H. A. (2008). Event-related potentials in the study of language. *Concise encyclopedia of brain and language*, 91-104.
Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and language*, *120*(2), 135-162.
Stowe, L. A., Kaan, E., Sabourin, L., & Taylor, R. C. (2018). The sentence wrap-up dogma. *Cognition*, *176*.
Stuellein, N., Radach, R. R., Jacobs, A. M., & Hofmann, M. J. (2016). No one way ticket from orthography to semantics in recognition memory: N400 and P200 effects of associations. *Brain research*, *1639*, 88-98.
Tanner, D., & Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia*, *56*, 289-301.
Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequencyinevent-related brainpotentials. *Memory & cognition*, *18*, 380-393.
Vigliocco, G., Butterworth, B., & Garrett, M. F. (1996). Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition*, *61*(3), 261-298.
Vikner, S. (2009). Agreement with predicative adjectives, with non-finite verbs and with predicate nominals. workshop on clausal and nominal parallels. University of Aarthus,

Vos, S. H., Gunter, T. C., Kolk, H. H., & Mulder, G. (2001). Working memory constraints on syntactic processing: An electrophysiological investigation. *Psychophysiology*, *38*(1), 41-63.
Walla, P., Endl, W., Lindinger, G., Lalouschek, W., Deecke, L., & Lang, W. (1999). Early occipito-parietal activity in a word recognition task: an EEG and MEG study. *Clinical neurophysiology*, *110*(8), 1378-1387.
Wechsler, S. (2009). Agreement features. *Language and linguistics compass*, *3*(1), 384-405.
Westergaard, M., Mitrofanova, N., Mykhaylyk, R., & Rodina, Y. (2017). Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism*, *21*(6), 666-682.
Yamada, Y., & Neville, H. J. (2007). An ERP study of syntactic processing in English and nonsense sentences. *Brain research*, *1130*, 167-180.

Appendices

Appendix 1: List of experimental stimuli

Festningen var gammel.

Utstyret var gammel.

Salden var gammel.

Saldet var gammel.

Prosessen var rask.

Toget var rask.

Berosken var rask.

Berosket var rask.

Skjorten var hvit.

Skapet var hvit.

Splyven var hvit.

Splyvet var hvit.

Drikken var varm.

Klimaet var varm.

Bistynen var varm.

Bistynet var varm.

Rommet var lyst.

Leiligheten var lyst.

Blåget var lyst.

Blågen var lyst.

Bildet var pent.

Fargen var pent.

Lergiet var pent.

Lergien var pent.

Interiøret var lekkert.

Fargen var lekkert.

Anskyltet var lekkert.

Anskylten var lekkert.

Huset var fint.

Filmen var fint.

Pedlevet var fint.

Pedleven var fint.

Temaet var enkelt.

Jobben var enkelt.

Pilaset var enkelt.

Pilasen var enkelt.

Fjellet var høyt.

Bygningen var høyt.

Optuget var høyt.

Optugen var høyt.

Leiligheten var ny.

Programmet var ny.

Låsken var ny.

Låsket var ny.

Rommet var tomt.

Flasken var tomt.

Toglet var tomt.

Toglen var tomt.

Bilen var gul.

Huset var gul.

Lanusken var gul.

Lanusket var gul.

Melken var sur.

Vannet var sur.

Lørsken var sur.

Lørsket var sur.

Butikken var dyr.

Kjøpet var dyr.

Maskomen var dyr.

Maskomet var dyr.

Smaken var god.

Resultatet var god.

Røglen var god.

Røglet var god.

Brødet var ferskt.

Frukten var ferskt.

Kjondlet var ferskt.

Kjondlen var ferskt.

Glasset var hvitt.

Boksen var hvitt.

Peliaget var hvitt.

Peliagen var hvitt.

Huset var rent.

Bilen var rent.

Glødset var rent.

Glødsen var rent.

Butikken var ny.

Bildet var ny.

Hinallen var ny.

Hinallet var ny.

Landet var vakkert.

Melodien var vakkert.

Skenaget var vakkert.

Skenagen var vakkert.

Rommet var kaldt.

Sommeren var kaldt.

Mutlet var kaldt.

Mutlen var kaldt.

Vitsen var morsom.

Innlegg var morsom.

Tagren var morsom.

Tagret var morsom.

Kaféen var åpen.

Rommet var åpen.

Vedlen var åpen.

Vedlet var åpen.

Været var kaldt.

Natten var kaldt.

Blalogtet var kaldt.

Blalogten var kaldt.

Skjorten var blå.

Havet var blå.

Luskinen var blå.

Luskinet var blå.

Himmelen var mørk.

Rommet var mørk.

Relset var mørk.

Relset var mørk.

Køen var lang.

Innlegget var lang.

Turanken var lang.

Turanket var lang.

Fargen var vakker.

Stedet var vakker.

Padullen var vakker.

Padullet var vakker.

Luften var ren.

Bordet var ren.

Delden var ren.

Deldet var ren.

Klimaet var tørt.

Vinen var tørt.

Såplet var tørt.

Såplen var tørt.

Vannet var skittent.

Snøen var skittent.

Temnet var skittent.

Temnen var skittent.

Plassen var åpen.

Brevet var åpen.

Stulten var åpen.

Stultet var åpen.

Biblioteket var stort.

Sengen var stort.

Bemunalet var stort.

Bemunalen var stort.

Gresset var vått.

Kluten var vått.

Lodlet var vått.

Lodlen var vått.

Sofaen var myk.

Smøret var myk.

Søklen var myk.

Søklet var myk.

Gutten var høy.

Treet var høy.

Våklet var høy.

Våklen var høy.

Boken var tykk.

Teppet var tykk.

Strilten var tykk.

Striltet var tykk.

Resultatet var positivt.

Responsen var positivt.

Smylget var positivt.

Smylgen var positivt.

Bildet var fint.

Sangen var fint.

Lunset var fint.

Lunsen var fint.

Fargen var grønn.

Gresset var grønn.

Tilnen var grønn.

Tilnet var grønn.

Gaven var dyr.

Utstyret var dyr.

Gulloten var dyr.

Gullotet var dyr.

Lyset var gult.

Kjolen var gult.

Klodliset var gult.

Klodlisen var gult.

Programmet var spesielt.

Saken var spesielt.

Slirget var spesielt.

Slirgen var spesielt.

Jakken var gul.

Eplet var gul.

Muljaet var gul.

Muljaen var gul.

Prosessen var langsom.

Tempoet er langsom.

Solutren var langsom.

Solutret var langsom.

Bygget var gammelt.

Kirken var gammelt.

Terroftet var gammelt.

Terroften var gammelt.

Miljøet var trygt.

Vein var trygt.

Jælset var trygt.

Jælsen var trygt.

Planen var enkel.

Eksempelet var enkel.

Reklasjunen var enkel.

Reklasjunet var enkel.

Terrenget var flatt.

Overflaten var flatt.

Jedget var flatt.

Jedgen var flatt.

Været var varmt.

Sommeren var varmt.

Histalyret var varmt.

Histalyren var varmt.

Sekken var stor.

Bordet var stor.

Folven var stor.

Folvet var stor.

Ideen var dum.

Spørsmålet var dum.

Gasken var dum.

Gasket var dum.

Historien var morsom.

Spillet var morsom.

Døkken var morsom.

Døkket var morsom.

Barnet var sykt.

Hunden var sykt.

Svarlyriet var sykt.

Svarlyrien var sykt.

Sengen var myk.

Teppet var myk.

Annloren var myk.

Annloret var myk.

Skjerfet var rødt.

Genseren var rødt.

Sjestøret var rødt.

Sjestøren var rødt.

Speilet var stort.

Sengen var stort.

Vanlet var stort.

Vanlen var stort.

Vannet var friskt.

Smaken var friskt.

Styplet var friskt.

Styplen var friskt.

Området var populært.

Serien var populært.

Doppstet var populært.

Doppsten var populært.

Blomstene var vakre

Blomsten var vakre

Starklene var vakre

Starklen var vakre

Bygningene var høye

Bygningen var høye

Fotrømmene var høye

Fotrømmen var høye

Spørsmålene var vanskelige

Spørsmålet var vanskelige

Dynkledene var vanskelige

Dynkledet var vanskelige

Husene var røde

Huset var røde

Nodlene var røde

Nodlet var røde

Bilene var hvite

Bilen var hvite

Tasmossene var hvite

Tasmossen var hvite

Rommene var romslige

Rommet var romslige

Slympene var romslige

Slympet var romslige

Bøker var interessante

Boken var interessante

Smaturene var interessante

Smaturen var interessante

Fruktene var friske

Frukten var friske

Gloskene var friske

Glosken var friske

Historiene var morsomme

Historien var morsomme

Fiavene var morsomme

Fiaven var morsomme

Byene var små

Byen var små

Spyrdene var små

Spyrden var små

Oppgavene var vanskelige

Oppgaven var vanskelige

Klivrene var vanskelige

Klivren var vanskelige

Vinduene var runde

Vinduet var runde

Plaflene var runde

Plaflet var runde

Resultatene var fantastiske

Resultatet var fantastiske

Boleggene var fantastiske

Bolegget var fantastiske

Prisene var lave

Prisen var lave

Eierduene var lave

Eienduen var lave

Nyhetene var triste

Nyheten var triste

Viglussene var triste

Viglussen var triste

Dagene var lange

Dagen var lange

Kjosperene var lange

Kjosperen var lange

Fargene var vakre

Fargen var vakre

Tjurpene var vakre

Tjurpen var vakre

Rommene var koselige

Rommet var koselige

Råmpelene var koselige

Råmpelet var koselige

Studentene var flinke

Studenten var flinke

Plimstene var flinke

Plimsten var flinke

Flaskene var tomme

Flasken var tomme

Styskene var tomme

Stysken var tomme

Dørene var åpne

Døren var åpne

Tareggene var åpne

Tareggen var åpne

Spillene var enkle

Spillet var enkle

Spukene var enkle

Spuket var enkle

Rettene var deilige

Retten var deilige

Folpatene var deilige

Folpaten var deilige

Bildene var flotte

Bildet var flotte

Juldene var flotte

Juldet var flotte

Kakene var søte

Kaken var søte

Bastene var søte

Basten var søte

Hundene var rolige

Hunden var rolige

Svurpene var rolige

Svurpen var rolige

Kundene var fornøyde

Kunden var fornøyde

Kluslene var fornøyde

Kluslen var fornøyde

Eplene var røde

Eplet var røde

Laplene var røde

Laplet var røde

Lysene var grønne

Lyset var grønne

Veplavene var grønne

Veplavet var grønne

Tallene var lave

Tallet var lave

Stjolksene var lave

Stjolksen var lave

Vinduene var store

Vinduet var store

Svukkene var store

Svukket var store

Rommene var rene

Rommet var rene

Histollene var rene

Histollet var rene

Spillene var populære

Spillet var populære

Avlaggene var populære

Avlagget var populære

Temaene var interessante

Temaet var interessante

Læmnene var interessante

Læmnet var interessante

Bildene var gamle

Bildet var gamle

Moslene var gamle

Moslet var gamle

Kakene var deilige

Kaken var deilige

Miftene var deilige

Miften var deilige

Posene var tunge

Posen var tunge

Plurfene var tunge

Plurfen var tunge

Produktene var unike

Produktet var unike

Blutene var unike

Blutet var unike

Billettene var billige

Billetten var billige

Slåpene var billige

Slåpen var billige

Guttene var nydelige

Gutten var nydelige

Sæplene var nydelige

Sæplen var nydelige

Husene var nye

Huset var nye

Hysmene var nye

Hysmet var nye

Dagene var varme
Dagen var varme
Slidene var varme
Sliden var varme
Putene var myke
Puten var myke
Peklatene var myke
Peklaten var myke
Fjellene var høye
Fjellet var høye
Dulvene var høye
Dulvet var høye
Jakkene var billige
Jakken var billige
Stjangene var billige
Stjangen var billige
Butikkene var åpne
Butikken var åpne
Hisklene var åpne
Hisklen var åpne
Sengene var komfortable
Sengen var komfortable
Katondene var komfortable
Katonden var komfortable
Barna var flinke
Barnet var flinke
Grumlene var flinke
Grumlet var flinke
Vintrene var kalde
Vinteren var kalde
Serakkene var kalde
Serakken var kalde
Eplene var gule
Eplet var gule

Stråsene var gule
Stråsket var gule
Barna var lykkelige
Barnet var lykkelige
Kjarvene var lykkelige
Kjarvet var lykkelige
Fagene var obligatoriske
Faget var obligatoriske
Graggene var obligatoriske
Gragget var obligatoriske
Skipene var store
Skipet var store
Sæplene var store
Sæplet var store
Gulvene var glatte
Gulvet var glatte
Bostruene var glatte
Bostruet var glatte
Skapene var nye
Skapet var nye
Kaslorvene var nye
Kaslorvet var nye
Kursene var lange
Kurset var lange
Trødlene var lange
Trødlet var lange
Bordene var små
Bordet var små
Kuvangene var små
Kuvanget var små
Rommene var ledige
Rommet var ledige
Nørlene var ledige
Nørlet var ledige

Gatene var brede

Gaten var brede

Galsoriene var brede

Galsorien var brede

Lysene var hvite

Lyset var hvite

Svadlene var hvite

Svadlet var hvite

Appendix 2: List of fillers:

Han spiste et grønt eple

Han spiste et grønn eple

Hun kjøpte en hvit kjole

Hun kjøpte en hvitt kjole

Det er en seriøs bedrift

Det et en seriøst bedrift

Hun leser en morsom bok

Hun leser en morsomt bok

Det var en slitsom tur

Det var en slitsomt tur

Det er en populær sang

Det er en populært sang

Huset har en lys stue

Huset har en lyst stue

Det var en enkel grunn

Det var en enkelt grunn

De tok et kult bilde

De tok et kul bilde

Dette var en seriøs diskusjon

Dette var en seriøst diskusjon

Han har en stor hund

Han har en stort hund

Det finnes en vakker park

Det finnes en vakkert park

Hvert hus har en unik design

Hvert hus har en unikt design

Dette er et sentralt tema

Dette er et sentral tema

Elevene snakket med en ung lærer

Elevene snakket med en ungt lærer

Hun kjøpte en ny sykkel

Hun kjøpte en nytt sykkel

Det finnes et stort maleri

Det finnes et stor maleri

Det finnes et lite museum

Det finnes et liten museum

Huset har et lyst kjøkken

Huset har et lys kjøkken

De snakker et lokalt språk

De snakker et lokal språk

Mange leter etter en pen bil

Mange leter etter en pent bil

Turister bor på et fint hotell

Turister bor på et fin hotell

Det var et godt påfunn

Det var et god påfunn

Han mottok et positivt svar

Han mottok et positiv svar

Mangfold er et stort tema

Mangfold er et stor tema

Historie er et sentralt fag

Historie er et sentral fag

Skolen et intensivt kurs

Skolen tilbyr et intensiv kurs

Studenter diskuterte et aktuelt emne

Studenter diskuterte et aktuell emne

Han skrev et vakkert dikt

Han skrev et vakker dikt

De fikk et langt brev

Det fikk et lang brev

Studenter trenger praktiske fag

Studenter trenger praktisk fag

Gutten lekte med små hunder

Gutten lekte med liten hunder

Butikken har rimelige priser

Butikken har rimelig priser

Mange liker lyse farger
Mange liker lys farger
Skoler innfører strenge regler
Skoler innfører streng regler
Noen sover med åpne vinduer
Noen sover med åpent vinduer
Han fant gamle brev hjemme
Han fant gammelt brev hjemme
Studenter hadde enkle spørsmål
Studenter hadde enkelt spørsmål
De lagde enkle regler
De lagde enkel regler
Hotellet har elegante rom
Hotellet har elegant rom
Damer liker fargerike blomster
Damer liker fargerik blomster
De kjøper dyre produkter
De kjøper dyrt produkter
Interiøret har lekre farger
Interiøret har lekker farger
Folk liker vakre sanger
Folk liker vakker sanger
De har runde bord
De har rundt bord
Politikere trekker fram viktige emner
Politikere trekker fram viktig emner
Hun fikk nydelige blomster
Hun fikk nydelig blomster
De hadde kreative ideer
De hadde kreativ ideer
Byen har bratte fjell
Byen har bratt fjell
Hundene trives med rolige dager
Hundene trives med rolige dager

Det utdannes flinke studenter

Det utdannes flink studenter

Barn liker varme drikker

Barn liker varm drikker

Folk liker ulike spill

Folk liker ulik spill

Glade barn lekte utenfor

Glad barn lekte utenfor

Studenter leste korte tekster

Studenter leste kort tekstene

Alle liker søte frukter

Alle liker søt frukter

Gamle bygg må bevares

Gammelt bygg må bevares

Turister besøker historiske steder

Turister besøker historisk steder

Trær blir dekket av grønne blader

Trær blir dekket av grønt blader

Bøker inneholder fargerike bilder

Bøker inneholder fargerikt bilder