Code-Switching in Bilingual Norwegian-English Speakers at the V2 Word Order Position

Annalisa Arcidiacono

Master's Thesis for the Degree of

Master of Philosophy in Linguistics

Institutt for lingvistiske, litterære og estetiske studier

Universitetet i Bergen

Norge/Norway

2022

Year:	2022
Title:	Code-Switching in Bilingual Norwegian-English Speakers at the V2 Word Order Position
Author:	Annalisa Arcidiacono

Acknowledgements

After a long and intense year, finally the day has arrived: writing these words of gratitude is the final touch of my thesis. It was a time of profound learning, not only academically, but also personally. Writing this thesis had a strong impact on my personality. I would like to say a few words to thank all the people who have supported and helped me during this period.

First of all, I would like to thank my supervisors, Tori Larsen and Christer Johansson, for all the help, support and guidance through the whole process of the thesis, and for helping me with the analysis of the data and the useful advice. Without them this study would not have come to life.

Further, I want to express my gratitude to the students who agreed to participate in my study, and the professors who let them.

A big thank you to my mother and father who, with their sweet and tireless support, allowed me to get here, contributing to my personal development. Thank you for your wise advice and your ability to listen to me. You have always been by my side, I love you endlessly.

Special thank you to My Lordship Louisa and Larsito, for many hours of help at Kamelåså. You have been an invaluable aid throughout the process. Thank you for your continuous and tireless support, without you I probably would not have made it. I love you mini Chomskys.

Louisa, I am so happy to call you my best friend, your pure soul always pushes me to do my best. You taught me so much, and I will never be able to express how grateful I am for it and for you.

Larsito, you were always there for me in the darkest of times. Thank you for all the nights spent in Casa Picazzo, watching stupid videos, top tier movies, listening to music and especially talking about life. You are the best.

Thank you to the rest of my family and friends for inspiration, love and moral support. All of you have had a decisive role in achieving this result, the point of arrival and at the same time the starting point of my life. Thank you for sharing the most important experiences with me over the years, I love you 3000.

Last but not least, an out of the ordinary ačiū to Sima La Bambi, with whom I have spent countless days and nights talking about very deep topics as well as very senseless ones. You have enriched and enlightened me like no one else has ever done. I could not have gone through these years and life in general without you, my irreplaceable best friend.

Abstract

This study overviews the major issues in the linguistic study of code-switching (CS), the alternating use of two languages in the same stretch of discourse by a bilingual speaker.
I investigated the effect of CS to L1 Norwegian and L2 English in V2 word order sentences at the point where the V2 requirement needs to be implemented in Norwegian, but not in English because of syntactic constraints. The aim of the research is to investigate if the reaction time in the different sentence conditions will be different at the test position of the sentences. Hence, the one containing the V2 word order in the L2 to L1 sentences (1) compared to the ones that do not need the V2 word order application, thus in the L1 to L2 sentences (2). According to the predictions, it would mean that from the testing position, the second segment of the sentence where the VP is presented, the phrase would be more difficult to process due to the code switching condition and the V2 requirement. Thus, the speakers would need more time to activate a different grammar and to apply it to the grammar of the language switched to. If it is not more difficult, then that could mean that both grammars are active in parallel. Based on the findings of past research (Schwochow et al., 2021), I hypothesize that the RTs will be faster when the participant will switch from L2 English to L1 Norwegian, because even if their L2 is strong, the L1 is still the dominant language system. Moreover, I also theorize that the sentences that contain both CS and V2 will be even more difficult to process, resulting in a higher RT.
The transfer of the grammatical property has been studied in L2 English learners with different L1s, but my study was mainly motivated by the work done by Johansson, Rugaard and Asperheim (2019). They investigated the effect of Code Switching (CS) to a second language (L2) in Garden Path (GP) sentences at the point of a possible GP, which had a follow up study conducted by Schwochow, Larsen and Johansson (2021). They predicted that code switching without V2 will be more accepted as it follows the predicted English word order, even when it conflicts with a grammatical rule in Norwegian. Their finding also predicts that CS involving a V2 word order change will be avoided and under-represented compared to code switching without V2. I have chosen to create the experiment according to the grammatical maze structure (Forster, 2010), where subjects were asked to select the grammatical continuation at the point of a possible V2. The GMaze was conducted for English to Norwegian (L1) with CS to L1 and with V2 word order (A), and for Norwegian to

English (L2) with CS to L2 and no V2 word order (B).

Code Switching:

English to Norwegian (A):

(A) Before eight o'clock [spiser han || han spiser] pannekaker og bacon til frokost.

Norwegian to English (B):

(B) Før klokka åtte [he eats|| eats he] pancakes and bacon for breakfast.

I have decided to work on Schwochow, Larsen and Johansson (2021) predictions, and my data support that both CS and V2 had longer reading times compared to a baseline without CS or V2 on the individual conditions. Two baselines were selected in order to isolate the effect of CS and V2. Significant differences from the baseline were detected, but there were no specific effects of which language it was code-switched to and from, only the general CS and V2 effects.

In conclusion, CS did add extra difficulty for reading or deciding on the testing sentences, but the data does not support a significant difference based on the factor of Language. A significant increase in RT was also detected when switching to L1 or L2 due to V2.

Abstrakt

Denne studien gir en oversikt over hovedproblemene i den språklige studien av kodebytte (CS) og den vekslende bruken av to språk i samme diskurs av en tospråklig taler.
Jeg undersøkte effekten av CS til L1 norsk og L2 engelsk i V2-ordrekkefølgesetninger der V2-kravet må implementeres på norsk, men ikke på engelsk på grunn av syntaktiske begrensninger. Målet med forskningen er å undersøke om reaksjonstiden i de forskjellige setningsforholdene vil være forskjellig ved testposisjonen til setningene. Den som inneholder V2-ordrekkefølgen i L2 til L1 setningene (1) sammenlignet med de som ikke trenger V2-ordrekkefølgeapplikasjon, altså i L1 til L2-setningene (2). I følge prediksjonen vil det bety at fra testposisjonen, det andre segmentet av setningen der VP presenteres, ville frasen være vanskeligere å behandle på grunn av kodebyttetilstanden og V2-kravet. Dermed vil deltagerene trenge mer tid på å aktivere en annen grammatikk og anvende den til språket som er byttet til. Hvis det ikke er vanskeligere, kan det bety at begge grammatikkene er aktive parallelt. Basert på funnene fra tidligere forskning (Schwochow et al., 2021), antar jeg at RT vil være raskere når deltakeren vil bytte fra L2 engelsk til L1 norsk, for selv om deres L2 er sterk, er L1 fortsatt det dominerende språksystemet. setningene som inneholder både CS og V2 vil mest sannsynlig være enda vanskeligere å behandle, noe som resulterer i en høyere RT. Overføringen av den grammatiske egenskapen har blitt studert hos L2 engelske elever med forskjellige L1er, men studien min var hovedsakelig motivert av arbeidet utført av Johansson, Rugaard og Asperheim (2019). De undersøkte effekten av kodebytte (CS) til et andrespråk (L2) i Garden Path (GP)-setninger ved punktet til en mulig GP, som hadde en oppfølgingsstudie utført av Schwochow, Larsen og Johansson (2021). De spådde at kodebytte uten V2 vil bli mer akseptert ettersom den følger den predikerte engelske ordrekkefølgen, selv når den er i konflikt med en grammatisk regel på norsk. Funnene deres forutsier også at CS som involverer en V2-ordrekkefølgeendring vil bli unngått og underrepresentert sammenlignet med kodebytte uten V2. GMaze ble utført for engelsk til norsk (L1) med CS til L1 og med V2-ordrekkefølge (A), og for norsk til engelsk (L2) med CS til L2 og ingen V2-ordrekkefølge (B).

Kodebytte:

Engelsk til norsk (A):

(A) Before eight o'clock [spiser han || han spiser] pannekaker og bacon til frokost.

Norsk til engelsk (B):

(B) Før klokka åtte [he eats|| eats he] pancakes and bacon for breakfast.

Jeg har bestemt meg for å jobbe med Schwochow, Larsen og Johansson (2021) forutsetningene. Mine data støtter at både CS og V2 hadde lengre lesetider sammenlignet med en baseline uten CS eller V2 på de individuelle forholdene. Jeg har valgt å lage eksperimentet i henhold til grammatical maze-strukturen (Forster, 2010), hvor forsøkspersonene ble bedt om å velge den grammatiske fortsettelsen ved punktet til en mulig V2. To grunnlinjer ble valgt for å isolere effekten av CS og V2. Signifikante forskjeller fra grunnlinjen ble oppdaget, men det var ingen spesifikke effekter av hvilket språk det ble kodesvitsjet til og fra, kun de generelle CS- og V2-effektene.

Som konklusjon la CS til ekstra vanskeligheter med å lese eller bestemme seg for testsetningene, men dataene støtter ikke en signifikant forskjell basert på språkfaktoren. En betydelig økning i RT ble også oppdaget ved bytte til L1 eller L2 på grunn av V2.

CONTENTS

Code-Switching in Bilingual Norwegian-English Subjects at the V2 Word Order Position.

1. Introduction

Psycholinguistics is the study of the mental representations and processes involved in language use, including the production, comprehension, and storage of spoken and written language. Its main focal areas have tended to be sentences and words, focusing on the generation of sentence structure and on syntactic planning, as well as on word finding and word building. Much of the study of comprehension has dealt with word recognition and sentence parsing in order to work out the syntactic structure of sentences (Warren, 2013). 'One question of interest for psycholinguists is the question of how closely real-time sentence processing routines align with grammatical knowledge: does the grammatical knowledge directly constrain sentence comprehension (and processing), or does it play a secondary role, 'cleaning up' the results of a comprehension process? (Townsend & Bever, 2001)?' Much experimental work has provided evidence for the view that the human sentence processor is constrained by grammatical knowledge even at the earliest stages of analysis, suggesting a very tight link between grammatical knowledge and the sentence processor (Dillon, 2009). To better understand this link, the difference between the sentence processor, also called the parser, which has to make decisions at the time words come in, and the grammar, which does not have such limitations, has to be pointed out. Indeed, the parser's task is different from the speaker's task. When we talk, we use grammar to build up sentences, but when we parse, we listen and do not only use grammar, but syntactic information and semantic correlations. However, a puzzle for this view is the observation that there are many apparently simple grammatical constraints that comprehenders seem unable to accurately apply during comprehension and processing (Wagers, Lau & Phillips, 2009). When we see language through such a framework it implies a strong distinction between a language system and how we understand and use it. This leads to an important theoretical question which is the focus of this study: if and why do bilingual and/or fluent speakers display a difference in reaction time during sentence comprehension of code switched sentences? In other words, does the grammar of L1 or L2 affect the comprehension of the other language system during code-switching? I look closer at code switching in Norwegian, which involves a V2 restriction. Both English and Norwegian sentences can start with an adverbial. But In Norwegian, because of the V2 word order requirement, a verb phrase is usually predicted, whereas in English, this predicts a subject noun phrase. If both grammars are active during the parsing of the initial adverbial, both predictions will be active and thus both continuations will be available (Schwochow, 2021).

The reason why I find it interesting to address this is that it provides insight about the cognitive process involved in language acquisition, and it may point towards implications for language teaching. The former notion is the main purpose of this thesis: to contribute to the current knowledge of the cognitive process of L2 acquisition.

In the first chapter relevant theoretical background will be discussed. The V2 word order, the topicalization of declarative sentences, bilingualism as well as second language acquisition will be introduced and discussed together with previous research. I also present the research questions and predictions the thesis is based on. The second chapter focuses on the ethical considerations taken into account for the current experiment. In the third chapter the research methodology is discussed, as well as the source of the data and the testing conditions. The fourth chapter is centered on the study's results and on the data analysis. The fifth chapter is reserved to the discussion of the results according to the research questions and predictions.

Finally, the sixth chapter is reserved to the conclusion. The seventh chapter offers the bibliography of the thesis, while the eight chapter is reserved for the appendices.

1.1 Theoretical Background

In this section of the thesis, I present relevant theoretical background for the current study: V2 word order, topicalization and code-switching, by introducing second language acquisition and bilingualism as well.

1.1.1 V2 Word Order

The V2 word order dictates that the verb must come in the second grammatical place in the sentence. For example, in English it would be 'I made breakfast', thus the sentence is structured in a SVO order. In Norwegian, the sentence follows the same structure: 'Jeg lagde frokost' ('I made breakfast'). But if the sentence starts with an adverbial phrase or an adverb like 'yesterday', in English it would be 'Yesterday, I made breakfast'. The V2 rule prohibits

Norwegian from following the same word order. The correct way to formulate that sentence in Norwegian is ‘I går laget jeg frokost’, that translates as ‘*Yesterday made I breakfast”. The subject and verb are reversed, to ensure the verb takes the second position in the sentence.

The V2 rule is not unique to Norwegian. In fact, it applies in almost all the Germanic languages such as Swedish, Danish, Icelandic, Faroese, German and Dutch. But it doesn't exist in English (Fischer et al, 2000).

Being the V2 patterns a syntactic phenomenon, they have certain environments where they can and cannot be accepted. Syntactically speaking, V2 requires a left-peripheral head, usually C, with an occupied specifier and paired with raising the highest verb-auxiliary to that head. The effects of V2 are caused by the left-peripheral head, which sets further requirements on a phrase that occupies the initial position, so that this phrase may always have specific characteristics (Urk, 2020). In order to fill this requirement, the verb moves. Unlike Norwegian, there is no V2 requirement in English. Instead, the lexical verb stays in the VP.

‘Modern English differs greatly in word order from other modern Germanic languages, but earlier English shared many similarities. For this reason, a description of Old English with V2 constraint as the norm has been proposed. The history of English syntax is thus seen as a process of losing the constraint’ (Fischer et al., 2000).

1.1.2 Topicalized Declarative Sentences

Topicalization is a mechanism of syntax that establishes an expression as the sentence or clause topic by having it appear at the front of the sentence or clause. This involves a phrasal movement of determiners, prepositions, and verbs to sentence-initial position (Sportiche et al., 2014).

In Norwegian, a topicalized declarative would have the word order AP-V-S-O (1). By moving the verb in front of the subject, the V2 rule is followed, even though an adverb has been placed in the initial position. Such a procedure is called inversion.

(1) I går lagde jeg frokost.

‘Yesterday I made breakfast’

In English, a topicalized declarative would have the word order AP-S-V-O (2). The verb phrase still follows the subject, and the sentence therefore keeps its SVO word order even though it has an adverb placed in the initial position.

(2) Yesterday I made breakfast.

While declarative sentences with an adverb in the initial position in Norwegian involve verb movement in front of the subject, the main verb does not move in the same construction in English.

1.1.3 Second Language Acquisition

Language transfer can be defined as "the influence resulting from the similarities and differences between the target language and any other language that has been previously (and perhaps imperfectly) acquired" (Odlin, 1989). It can facilitate language learning when the native language and target language have structures that are similar, such as positive transfer. When the native and target language have structures that differ, the learner is likely to make errors. This is called negative transfer. Transfer can affect both comprehension and production and affects all linguistic subsystems such as syntax, morphology, vocabulary and phonology.

1.1.4 Bilingualism and Code-Switching

CS is a linguistic behavior used among bilinguals for a variety of reasons. CS does not occur randomly, but rather is constrained by grammatical rules as well as social and cultural norms.
Zentella (1997) noted that adults tend to code-switch more in familiar informal settings where the conversation partners share the same languages. In more formal settings, however, adults tend to code-switch less.
There is a general consensus in the CS literature, that switches of the intrasentential type do not include randomly inserted elements in either language, but rather CS occurs fluently at

predictable syntactic points in the sentence and that the most proficient bilinguals are able to easily switch between languages within a sentence without violating the grammatical rules of either language (Bullock & Toribio, 2009). Bilinguals may activate their grammars in parallel, and they may have two predictions when their grammars predict alternative structures, as argued by Putnam et al. (2018).

1.1.5 Previous Research

The transfer of the grammatical property has been studied in L2 English learners with different L1s. It is worth mentioning that I will only be discussing studies concerning Norwegian L1.

1.1.5.1 Westergaard

Westergaard (2003) investigated word order acquisition in L1 Norwegian L2 English learners. The participants were students in grades 2- 7. Westergaard tested different structures, but the one I am more interested in is the topicalized structure. As already mentioned, the word order in these constructions differs in Norwegian and English, since Norwegian but not English has the V2 rule in topicalized declarative sentences. The results from the Westergaard study revealed that the participants in all ages showed “massive transfer of V2 word order” (2003). In topicalization structures among the 5th graders, 70% chose the V2 word order, and when presented with sentence pairs, they chose the ungrammatical V2 option. The 7th graders produced non-target structures 25% of the time. The older the participants, the fewer V2 errors were created in topicalised structures in English.

1.1.5.2 Johansson, Rugaard and Asperheim

Johansson, Rugaard and Asperheim (2019) investigated the effect of Code Switching (CS) to a second language (L2) in Garden Path (GP) sentences at the point of a possible GP, through Self Paced Reading (SPR). Significant effects were detected for GP, but no significant effect for CS and no additional time for the interaction of GP with CS. In conclusion, CS to L2 did not add extra difficulty for reading or deciding GP sentences. Thus switching from L1 to L2 has a low cost for bilinguals. The results indicate that CS may be beneficial as the cost of CS could be compensated by faster responses for resolving GP (Johansson et al., 2019).

1.1.5.3 Schwochow, Larsen and Johansson

Schwochow, Larsen and Johansson (2021) experiment was based on the previous work done by Johansson, Rugaard and Asperhein (2019). This time, the results suggested a faster processing time when switching to a L1. Also, the GP effects were strong. However, the interesting result is that code switching with V2 can have an effect that is as large as the effect of a GP, which is typically explained by a reparse. They predicted that code switching without V2 will be more accepted as it follows the predicted English word order, even when it conflicts with a grammatical rule in Norwegian. Their finding also predicts that CS involving a V2 word order change will be avoided and under-represented compared to code switching without V2. The latter is one main reason as to why my experiment is based on their experiment, but not including the GP condition.

1.2 Research Questions and Hypotheses

1.2.1 Research Questions

As already mentioned, the aim of this study is to examine code-switched sentences at the V2 word order position in bilingual speakers of Norwegian (system that requires a V2 word order) and English (that does not require a V2 word order) in adults without any communication disorders when processing written input of language. Furthermore, the study aims to investigate:

Research Question 1: Is there a difference in RT from L1 to L2 or vice versa when processing the code switched sentences?

Research Question 2: Is there a difference in RT when the participant is faced with the choice at the testing position to either select the correct grammar for the starting language, but incorrect for the ending language, or to select the correct grammar for the for the ending language (correct choice in this experiment) but incorrect for the starting language in a code-switching environment that contain also the V2 requirement in L1 Norwegian?

In order to accomplish the aim of the study, a G-maze experiment was created. To explain my hypotheses, an example of a test sentence is introduced, to then be better explained in the method section. Sentence (3) is a L2 English to L1 Norwegian code-switched test sentence that also requires the V2 word order selected because the syntactical structure in Norwegian demands the verb to be in the second position. Sentence (4) is a L1 Norwegian to L2 English code-switched test sentence that does not demand the V2 word order, since English does not have this requirement.

<u>Code Switching</u>:

English to Norwegian (3):

(3) Before eight o'clock [spiser han || han spiser] pannekaker og bacon til frokost.

Norwegian to English (4):

(4) Før klokka åtte [he eats|| eats he] pancakes and bacon for breakfast.

1.2.2 Hypotheses

H0: The null hypothesis is that there is no variation between the selected baselines and the test conditions.

H1: My first alternative hypothesis is that the code-switching demands greater processing, thus a longer reaction time. The RTs will be faster when the participant will switch from L2 English to L1 Norwegian, because even if their L2 is strong, the L1 is still the dominant language system.

H2: My second alternative hypothesis is that the V2 word order demands greater processing, and also a grammar shift. Thus, the sentences that contain both CS and V2 will be even more difficult to process, resulting in a higher RT. If CS needs to activate a second grammar, this can be predicted to make a reparse involving cross-linguistic input combinatorially more difficult and thus adding CS in a V2 would lead to slower reactions.

H3: My third and last alternative hypothesis, additionally investigates the learning effect and the presentation order of the sentences in the experiment. The implications of the learning effect is that the participants get faster towards the end of the experiment, and it is not related to the sentence conditions.

These predictions are based on previous findings in research on transfer of the grammatical property in L2 English learners with different L1s such as Johansson, Rugaard and Asperheim (2019) study on the effect of code-switching to a second language in garden path sentences. The other study my hypotheses are mainly based on is Schwochow, Larsen and Johansson (2021) experiment on code-switching in garden path sentences that also involved V2 word order in certain conditions, that were only analyzed post hoc.

2. Ethical Considerations

2.1 'Do No Harm'

Bryman and Bell (2007) formulated ten principles related to ethical considerations in research. The most important ethical consideration is that research participants should not be subjected to harm. In simpler terms: "Do no harm", that is why it is important to predict who could access the data and the possible future use of the study. In this case, only the researchers can access the raw data, but what is yet to come needs to be kept in mind. For this reason, I let the participants know through a participant information and consent form (Appendix B) that "I will use the reaction time data obtained from your participation solely for the purpose of this particular study. The results of this study will be used in publications and presentations, including a master thesis and possible PhD dissertation". Thus, I guaranteed that their data will not be divulged to third parties, which will have access only to the report.

2.2 Transparency

Any type of communication in relation to the research should be done with honesty and transparency. Indeed, my aim was to make clear what my interest was, and what my research will be used for. However, a remark should be made: I did not intend to lie, but I needed to give limited information about the study itself to the participants at the beginning of each test, otherwise the results could be distorted. That is the main reason why, when looking for participants, I only mentioned that I was currently working on a psycholinguistic experiment for my Master thesis, a lexical decision task on code switching in bilingual speakers. No other information was given about the experiment. I did offer a debriefing at the end of each session, in which I asked for feedback and I explained and discussed the topic.

2.3 Informed Consent

According to Saunders, Lewis and Thornhill (2012) “the principle of informed consent involves researchers providing sufficient information and assurances about taking part to allow individuals to understand the implications of participation and to reach a fully informed, considered and freely given decision about whether or not to do so, without the exercise of any pressure or coercion”. For this reason, at the beginning of the experiment, participants give their consent with the possibility of withdrawal from the study at any time, as follows: “participation in the study is voluntary. If you choose not to participate, you can withdraw your consent at any time without giving a reason. There will be no negative consequences for you if you choose not to participate”. Subjects are tested individually and are required to read instructions (Appendix C) on a computer screen. By accepting to continue after the information about the task is given, besides reading and signing a consent form (Appendix B), subjects document that they understand the instructions and consent to participate.

2.4 Privacy and Anonymity

Moreover, the protection of the privacy of research participants has to be ensured as well as anonymity of individuals and organizations participating in the research, for this reason each participant is assigned a number, from one to twenty-five. I did ask for personal information such as gender, age and education, as well as if the participants are native speakers without any communication disorder. After pondering whether to ask for this information or not, I have come to the conclusion that it might be relevant in the future to know their gender education and age for further analysis, compare different grades of education, as well as gender and different age groups.

2.5 Inclusion

Respect for the dignity of research participants should be prioritized, and it relates to who to include in the research, and possible stigmatization of the participants. If I am to find out that a participant has some disability that could affect the results, even if I am not in the position to give any medical diagnosis, I would let the participant continue the test, but I would indicate the test as a possible outlier and I might not include it in the final results of data analysis. Only researchers are able to detect which participant is who, but this information is not stored with the data.

2.6 Compensation

Furthermore, the participants are required to be in a room for an average of fifteen minutes, and as shown in the "Participant Information Sheet" (Appendix B), there is no compensation for their participation, mainly because there could be a conflict: the participants could hesitate to leave the experiment if offered compensation. On the other hand, if there is no reward, there would not be any motivation. For this reason, when asking to join the experiment, I stated that they were going to volunteer their time.

3. Research Methodology

3.1 Type of Research and Rationale

The overall approach to the research is quantitative (Creswell, 2009), which is an approach for testing objective theories by examining the relationship among variables, as I aimed to produce generalizable knowledge about the causes of a phenomenon. These variables, in turn, can be measured, typically on instruments, so that numbered data can be analyzed using statistical procedures. The type of data is quantitative, since I have collected numeric data, the RTs. Moreover, the design is experimental, and seeks to determine if a specific treatment influences an outcome, thus I did gather experimental data by controlling and manipulating variables.

Based on the conditions L2 (English) to L1 (Norwegian) code-switched sentences compared to L1 (Norwegian) to L2 (English) code-switched sentences, and monolingual English sentences compared to Norwegian monolingual sentences, that have been used as baseline sentences, I intended to discover:

A. If the RT that measures the time it takes to process either NOR to ENG or ENG to NOR CS sentences will be different from L2 to L1 sentences than from L1 to L2 sentences.

B. If the monolingual L2, the English one, would be slower than the monolingual L1, Norwegian.

Thus, I want to study if the RT in the different sentence conditions will be different at the test position of the sentences, the one containing the V2 word order in the L2 to L1 sentences compared to the ones that do not need the V2 word order application, thus in the L1 to L2 sentences.

3.1.1 Participants, Material and Tools

The participants (Table 1) in this study consisted of 25 adult bilingual speakers of Norwegian and English, sampled from around the HF building at University of Bergen and also through the help of Professor Vadim Kimmelman, Professor Koenraad de Smedt and Professor Carl Börstell, who advertised my need of participants with their students. They were tested to investigate the processing of written stimuli regarding code-switching and V2 word order.

GENDER	AGE	EDUCATION
Male: 10	18-25: 13	HS: 5
Female: 12	26-35: 11	BS: 10
Other: 3	36-45: 1	MS: 7
		Other: 2
		NR: 1

Table 1. This table displays the descriptive statistics for all 25 participants regarding gender, age and education. The NR responses mean that the participant(s) declined to provide an answer. While HS stands for Highschool degree, BS stands for Bachelor degree and MS stands for Master degree.

I have chosen to create the experiment according to the G-maze structure (Forster, 2010), which uses real word distractors that are anomalous given the context. It is a variant of the Self-Paced Reading (SPR) task, in which a series of sentence stimuli in smaller segments are presented, phrase by phrase. The participants determine the length of reading time, by choosing between grammatically possible and grammatically impossible phrases/sequences of words, for a specific verbal phrase (VP). Thus, each sentence starts off masked and the participant presses a button to reveal each successive word and mask the previous word, with the time between button presses constituting the word's reading time (RT) which is recorded and related to controlled positions in the sentence. I opted for a non-cumulative G-maze, in which the sequence of words disappear after being chosen, mainly because I wanted to avoid the occurrence of learning and memory strategies, and this method forces participants to read sentences sequentially, with no looking ahead or looking back. Moreover, the G-Maze had a centered display of the sentences on the screen but the parts of the sentence were linearly displayed in order to follow a more natural presentation.

The maze task does not require special equipment: all it needs is a way of displaying stimuli, the computer lab situated in a sound-prof room, and to record button-presses through the response pad.

It is worth mentioning that the SPR method has some disadvantages, such as the request to fulfill two tasks, comprehension and lexical decision, and also the need to use memory strategies. Moreover, since it is the participant that has to push the button on the response pad to go to the next phrase, the reaction and the reading time can vary, due to subjects and items but also because the participant gets tired and less focused. Thus, a longer reaction time, and therefore a high-processing difficulty can be due to the ambiguity, anomaly, or even to long-distance dependency. But this task measures the time each participant takes to recognize and assimilate each successive phrase in a sentence. Thus, it is possible to measure lexical access time during sentence processing, and the reaction time for each interested stimulus (that includes the time needed to make a decision). In order to design an experiment that would allow me to present stimuli and record responses, I have used the program "Superlab" (Heller et al., 2015).

I intended to investigate if the reaction time in the different sentence conditions would be different, at the verb position either with or without application of the V2 word order. According to the predictions, it would mean that from the testing position, the second segment of the sentence where the VP is presented, the phrase would be more difficult to process due to the code switching condition and the V2 requirement. Thus, the speakers would need more time to activate a different grammar and to apply it to the grammar of the language switched to.

3.1.2 Experimental Design and Procedure

The participants are instructed to read each phrase of the experiment carefully, make quick intuitive decisions, and to select one of the alternatives through the response pad. Subjects are tested individually and are required to read instructions (Appendix C) on a computer screen, and to answer questions about their gender, age and education. By accepting to continue after the information about the task is given, subjects document that they understand the instructions and consent to participate, besides reading and signing a consent form (Appendix B).

The maze task that I used for the experiment, presents sentence stimuli phrase-by-phrase, the first part and third part of each sentence stimulus does not require a decision, hence the participants are instructed to press the middle button to go forward with the experiment. The only sequence of frames with two alternatives, the second (middle) part of the sentence, presents one choice on the left side and one choice on the right side, and that is when the participants are required to press the left red button or right green button through a response pad. The procedure of how the sentences are presented phrase by phrase is shown in the flow chart diagram (Figure 1). The correct grammatical continuation of the sentence can be displayed either on the left or on the right (I made sure that there was an equal amount of right and left correct answers). The correct choice is underscored in the following example, which is an English to Norwegian CS sentence with the correct answer presented on the left (5). Each testing sentence is divided in three different segments, the first one formed by the topicalized clause; the second one which is the test segment, where the participants are required to take a decision; finally, the third segment is the rest of the sentence. The first and third segments are not taken into consideration for the analysis:

(5) Before eight o'clock [<u>spiser han</u> || han spiser] pannekaker og bacon til frokost.
'Before eight o'clock he eats pancakes and bacon for breakfast'

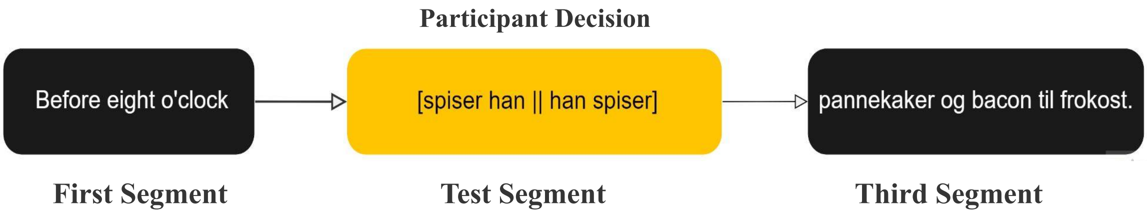

Figure 1. This flow chart diagram is a visual representation of how the sentences are presented phrase by phrase.

The ungrammatical, incorrect alternative of the sentence continuation is either the V2 word order wrongly applied to the English syntactic structure, or the lack of application of the V2 word order in the Norwegian syntactic structure. This is shown in (5), where the wrong choice presented on the right "han spiser" is grammatically wrong in Norwegian since it follows the English syntactic word order even after an adverbial phrase.

The training sentences have a different structure from the test sentences, they do not start with an adverbial phrase, and there is no need for V2 word order application. An example follows (6):

(6) Norge har en lang felles [with border || border with] Sweden.
'Norway has a long common border with Sweden'

The filler sentences also have a different structure, this helps to keep the participant from forming a strategy based on always seeing the same sentence structure. In addition, filler sentences are presented in a new randomly generated order for each participant, once ENG to NOR (7) and once NOR to ENG (8).

(7) The professor introduced a [nytt tema || tema nytt] for studentene.
'The professor introduced a new topic for the students'
(8) Professoren introduserte et [new topic || topic new] for the students.
'The professor introduced a new topic for the students'

Moreover, I have decided to use filler sentences as baseline sentences in order to isolate V2 through differences in RTs, since it cancels out the CS effect. I have discussed with both of my supervisors the possibility of using the filler sentences in the analysis prior to starting the experiment; however, I have decided to use the monolingual baseline. After drawing up the detailed analysis plan, I realized that I did need to use the filler sentences as baseline in order to isolate the V2 confusion effect, instead. Thus, this experiment has two different analyses: one in which the test sentences are compared to the filler baseline sentences in order to isolate the CS effect. In the second one, the test sentences are compared to the monolingual baseline sentences, in order to isolate the V2 effect. A more detailed explanation is given in the analysis section.

3.2 Source of the Data

In order to create the data set, I have decided that each of the 17 testing sentences would start with a different adverbial phrase in Norwegian. The correct answers were presented on the left and right in each condition and an equal number of times, in order to keep the participant from expecting the correct answer always in one position. I have then translated the sentences in English, for a total of 102 testing sentences, see (13) to (16) for examples.

The Baseline sentences are the monolingual sentences, both in English (9) and Norwegian (10), for a total of 34 sentences, and the Filler sentences, 8 in English to Norwegian (11) and 8 in Norwegian to English (12), for a total of 16 sentences. As already mentioned, the CS sentences English to Norwegian and Norwegian to English are presented twice, one time with the ‘correct’ answer on the right and one time with the ‘correct’ answer on the left, for a total of 68 sentences. A scheme with the structure and an example template of the test sentences follows:

<u>Baseline Monolingual</u>:

English (9):

(9) Before eight o'clock he eats pancakes and bacon for breakfast.

Norwegian (10):

(10) Før klokka åtte spiser han pannekaker og bacon til frokost.

'Before eight o'clock he eats pancakes and bacon for breakfast.'

Baseline Filler:

English to Norwegian (11):

(11) The professor introduced [et nytt tema || et tema nytt] for studentene.

'The professor introduced a new topic for the students.'

Norwegian to English (12):

(12) Professoren introduserte [a topic new || a new topic] for the students.

'The professor introduced a new topic for the students.'

Code Switching:

English to Norwegian Left (13):

(13) Before eight o'clock [spiser han || han spiser] pannekaker og bacon til frokost.

'Before eight o'clock he eats pancakes and bacon for breakfast.'

English to Norwegian Right (14):

(14) Before eight o'clock [han spiser || spiser han] pannekaker og bacon til frokost.

'Before eight o'clock he eats pancakes and bacon for breakfast.'

Norwegian to English Left (15):

(15) Før klokka åtte [he eats|| eats he] pancakes and bacon for breakfast.

'Before eight o'clock he eats pancakes and bacon for breakfast.'

Norwegian to English Right (16):

(16) Før klokka åtte [eats he || he eats] pancakes and bacon for breakfast.

'Before eight o'clock he eats pancakes and bacon for breakfast.'

As already mentioned, the experiment contains three separate test sections: the first one with the training sentences, the second containing the testing sentences and the third one containing the filler sentences. Both training and filler sections are made up of sentences designed with a similar length, but structurally dissimilar to the test sentences. The training sentences were included to make sure that the participants would understand the task, since I have also inserted a feedback, and a smiley emoji face in case they got the answer right and a sad emoji face in case they got the answer wrong. This helps the participants to learn how to complete the task. The filler sentences help to prevent them from understanding the pattern of the task, since the pattern is broken, and to avoid the occurrence of learning strategies.

I chose to use different adverbials to begin each sentence stimuli in order to add variation to them, this difference itself was not used as a factor in my experiment. Using different adverbs or adverbial phrases may or may not have an impact on the sentence processing and reaction times. Thus, this was taken into account using "Sentence" as a factor in my analysis.

It is worth mentioning that while I was creating the dataset, I found it more fitting to use mostly pronouns instead of NPs, especially at the V2 testing position. I wanted to control for the "heaviness" of the subject NP in the interesting clause, since pronouns and NPs might be resolved in different ways, thus triggering either grammar or mental lexicon for resolution (Nicol & Swinney, 1989).

3.3 Conditions

Before the explanation of the conditions, the idea of minimal comparison should be introduced. A minimal comparison is the idea that just one element changes, thus there is just one small change to be measured in each sentence. Thus, we take an element that we are not interested in, for example a baseline sentence, and then we add one small change, the element we are interested in, such as the code switched phrase that presents a V2 word order. Afterwards, we subtract the element we are not interested in, and from that reaction time, we are left with just the effect we are interested in. The G-maze contains an introduction with the explanation of the task, questions about gender, age and education of the participant and three separate test sections follow (randomized order per participant). I tested seventeen template sentences and each template contains six sentences (for a total of 102 sentences), as follows:

1. 6 Training Sentences.
2. 17 Baseline Monolingual English Sentences (8).
3. 17 Baseline Monolingual Norwegian Sentences (9).
4. 34 English to Norwegian CS Test sentences (12),(13).
5. 34 Norwegian to English CS Test sentences (14), (15).
6. 8 Baseline Filler Sentences English to Norwegian (10).
7. 8 Baseline Filler Sentences Norwegian to English (11).

Thus, each sentence template is formed by two baseline monolingual sentences, one fully in Norwegian and one fully in English. Two baseline filler sentences, one version from L1 to L2 and a second version from L2 to L1. Two English to Norwegian code-switched sentences, and two Norwegian to English code-switched sentences, both with one correct answer on the left and one with the correct answer on the right.

I have included different factors, both fixed and random, that I wanted to take into account in the experiment, some of them contain different levels. These factors were tagged in the data set and each data point was labeled with the corresponding level of said factor.

The fixed factors are:

1. Age.
2. Gender.
3. Education.
4. Language: Norwegian to English and English to Norwegian.
5. Presence of V2: Yes, no or NA. If the second part of the sentence is in English, there wouldn't be presence of V2, viceversa if the second part of the sentence is in Norwegian, there would be presence of V2.
6. Right or Left: Right, left or NA, tagging which choice was the correct one, also in order to have a balanced number of correct responses between the right side and left side.
7. Order, which was not a part of the original design, and was added later. It takes into account the learning effect, and it stands for the order and number of presentations,

including questions regarding age, gender and education, as well as training, filler and testing sentences.

The random factors are:

1. Sentence: From 1 until 17.
2. Participant, for a total of 25.

4. Statistical Analysis and Results

4.1 RT and Participants Outlier Analysis for Both Datasets

For the purpose of making the baseline, I first calculated the means, in order to get the RT Difference.

To evaluate the results, the data was exported from the program SuperLab. Some of the data was evaluated manually on Excel to remove outliers such as participant 24 (P24), since it has the highest number of errors, 19. The next highest is 7 errors, this shows a large difference between error rates and the reason why I have decided to remove P24 from the analysis, after I have looked at the error analysis in R. P24 was removed mainly because of the error rate and not because of the RT, since the RT is not overrepresented in the outliers. The high error rate is probably due to the fact that both answers are technically correct and the participant chose another preference and grammar, thus they might solve the task in a different way. Overall, P24 still got 93.8% (287/306) of responses correct. After the removal of the outlier participant, I set the RT boundaries to exclude high and low RT outliers. I had to create different datasets to isolate the RT, and to calculate the Mean and the Standard Deviation.

After some calculations, I have obtained the following (Table 2), where 'Upper B.' stands for upper boundary, and 'Lower B.' stands for lower boundary.

Condition	Mean	Upper B.	Lower B.
NORENG:	1702.5	3112.4	292.6
ENGNOR:	1680.3	3017.0	343.5

Table 2. This table shows the Mean, Upper and Lower Boundary for the two Language conditions in the code-switched sentences: NORENG that stands for a sentence that is being code switched from L1 Norwegian to L2 English. Similarly, ENGNOR stands for a sentence that starts with L2 English and code switched to a L1 Norwegian.

The main reason why I have decided to include only one decimal is because the model estimates down to 4 decimal points, but it is not possible to assume the model can actually estimate to that level of accuracy.

In the English Monolingual I calculated the following (Table 3), where Mean 1, 2 and 3 respectively stand for Mean of the first, second and third segment of the sentence, and their respective upper and lower limits.

	Means	Upper B.	Lower B.
Mean1	2150.3	4125.8	174.8
Mean2	3596.6	6121.8	1071.5
Mean3	4415.5	7547.3	1283.7

Table 3. This table shows the Means, Upper and Lower Limits of the three segments (Mean 1, 2 and 3) of the English Monolingual sentences.

In the Norwegian Monolingual we calculated the following (Table 4), where Mean 1, 2 and 3 respectively stand for Mean of the first, second and third segment of the sentence, and and their respective upper and lower limits:

	Means	Upper B.	Lower B.
Mean1	2007.8	3924.7	90.8
Mean2	3448.0	5923.6	972.4
Mean3	4282.7	7263.5	1301.9

Table 4. This table shows the Means, Upper and Lower Limits of the three segments (Mean 1, 2 and 3) of the Norwegian Monolingual sentences.

I have then created a single dataset that includes all the information from the just mentioned datasets, in order to work on RStudio (R Core Team, 2021) where I imported the final dataset and transformed numbers into numeric and columns/character into factors. R is a programming language used for statistical computing while RStudio, a graphical interface to R, uses the R language to develop statistical programs, and it works as an integrated development environment. The software package R is a powerful tool for performing various types of statistical analyses.

Later, I performed an outlier analysis of both participants and sentences, visible in the following assoc plot “Overall Error” (Figure 2), to analyze the Proportion of Errors per Participant. As already mentioned, P24 is the only participant that was excluded from the rest of the analysis and marked as an outlier because its error rate is the highest. It is possible to see that it differs significantly as illustrated in the assoc plot (Figure 2).

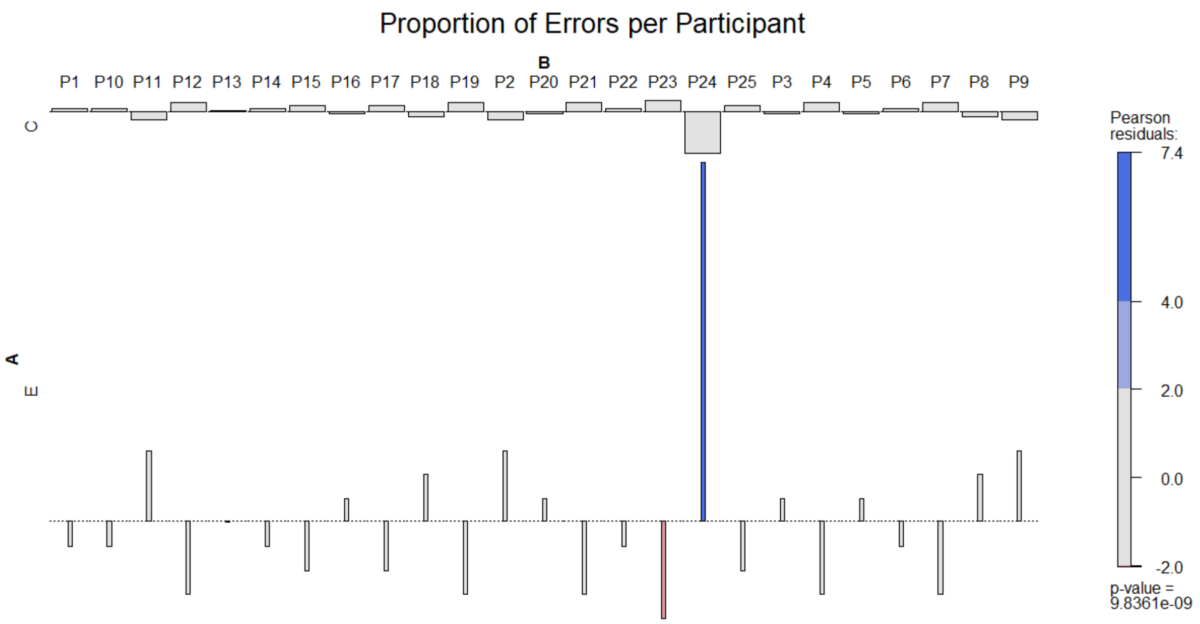

Figure 2. This association plot is a visual representation of the outlier analysis of both sentences and participants.

4.2 Baseline for the Monolingual and Filler Dataset

In order to analyze the data, I have created two datasets with two different baselines to isolate the CS and the V2 effect. The conditions that I have compared in the first dataset as a mean to isolate the specific effect of the CS (Monolingual sentences - code-switching = pure CS effect, referred to as 'difference') and get rid of the "noise" RT are:

1. Monolingual English vs CS English to Norwegian sentences
   English to Norwegian: EngNor + CS + V2

   CS: MonoNO – 0CS + V2

   +CS (0)V2
2. Monolingual Norwegian vs CS Norwegian to English sentences
   Norwegian to English: NorEng + CS + V2

   CS: MonoEN – 0CS + V2
   +CS (0)V2

The conditions that I have compared in order to isolate the specific effect of the V2 (filler sentences - V2= pure V2 effect, referred to as 'difference') and get rid of the "noise" are:

1. Filler English vs V2 English to Norwegian sentences
   EngNor + CS + V2

   FillerEngNor + CS + V2(0)

   (0)CS V2
2. Filler Norwegian vs V2 Norwegian to English sentences

   NorEng + CS + V2

   FillerNorEng + CS + V2(0)

   (0)CS V2

By 'noise' RT, I allude to the time that it takes to read, process the sentence, and take a decision, to switch to V2, and also to press the button on the response pad. The difference in RTs between baseline monolingual (V2) and baseline filler (CS) gives us a purer reaction time for CS and V2. The mean difference measures the absolute difference in RT between the mean value in two different groups.

4.3 Filler Baseline to Isolate V2

The distribution of the difference in RTs in the Filler Baseline Sentences is shown in the histogram (Figure 3), in order to bring out the pattern in the grouped frequency, and was created with the following formula:

```
hist <-hist(FillerBL$Diff.Filler, main="Histogram for Data from all Participants",
xlab="Difference in RT from Baseline", breaks = 50)
```

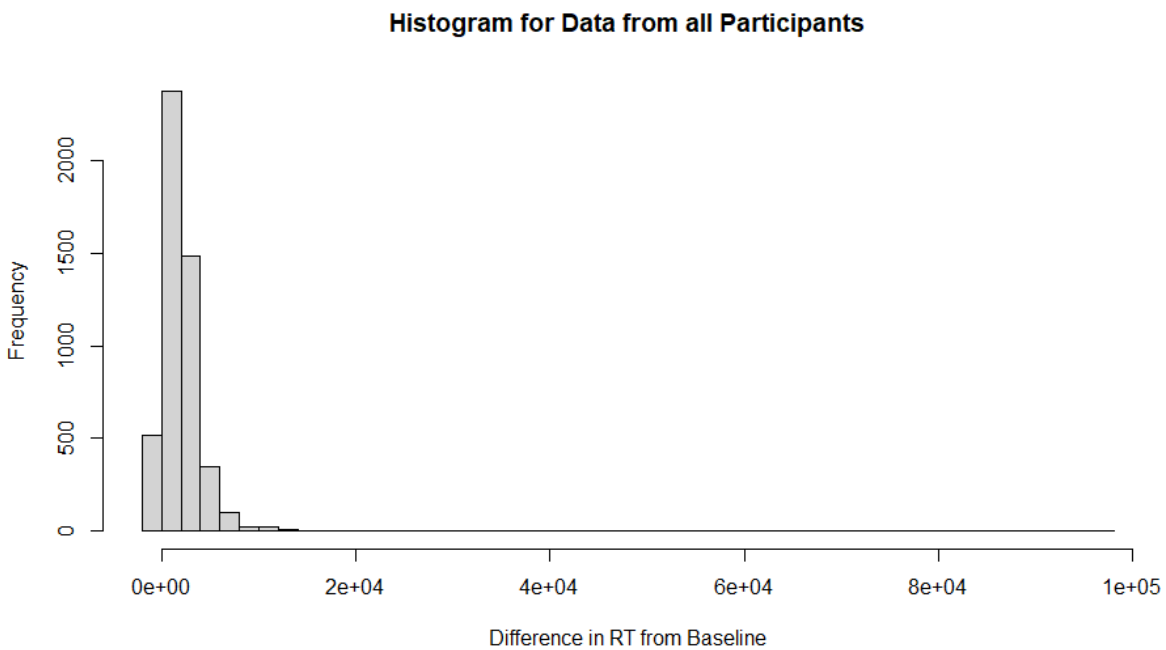

Figure 3. This histogram displays the distribution of the difference in RTs in the Filler Baseline sentences.

From the summary of the histogram in RStudio, I noticed that the values in the bins (columns in the histogram), had three very high data points, making the x axis stretch out. Thus, to focus on the test position, I picked out only the second segment of the filler sentences, the segment where the participants need to make a decision. Later, as already mentioned in the earlier section, in order to make the histogram less stretched out, I removed extreme outliers points (not regular ones) for “Difference” and extreme outliers RTs due to distractions or other factors, by evening out the histogram’s ends calculating the “Mean” of the “Difference” column in the dataset.

Data points in a normal distribution are more likely to fall closer to the mean. In fact, 68% of all data points will be within ±1SD from the mean, 95% of all data points will be within ± 2SD from the mean, and 99% of all data points will be within ±3SD. Statisticians have determined that values no greater than plus or minus 2 SD represent measurements that are closer to the true value than those that fall in the area greater than ± 2SD (Berger and Koefer, 2021). This is the main reason why I have decided to set the normal limits within ± 2SD. Also, as it is possible to see in (Figure 4) and (Figure 7), the histograms kept respectively 98.7% and 98.6% of the data, thus the majority of it. And since removing the outliers kept the majority of the data, there was no need to use a logarithmic transformation, that transforms the data in a logarithmic scale, compressing the data, so the tail was not abnormally long.

The two SD of the values in the “Difference” Column is “5578.7”. Thus, I added the two “Standard Deviation” to the “Mean” (The mean is 2280.0) = (7858.7), to obtain the upper limit. I have then removed the 2 SD from the “Mean” (-3298.6) in order to get the lower limit. I then kept all “Difference” values that are within two SD from the mean. The result is a Filler Baseline Histogram (Figure 4), only for the test position, using the following formula:

```
histfillerseg2 <-hist(FillerBLseg2$Diff.Filler, main="Histogram for Test Segment from all Participants", xlab="Difference in RT from Filler Baseline", breaks = 50)
```

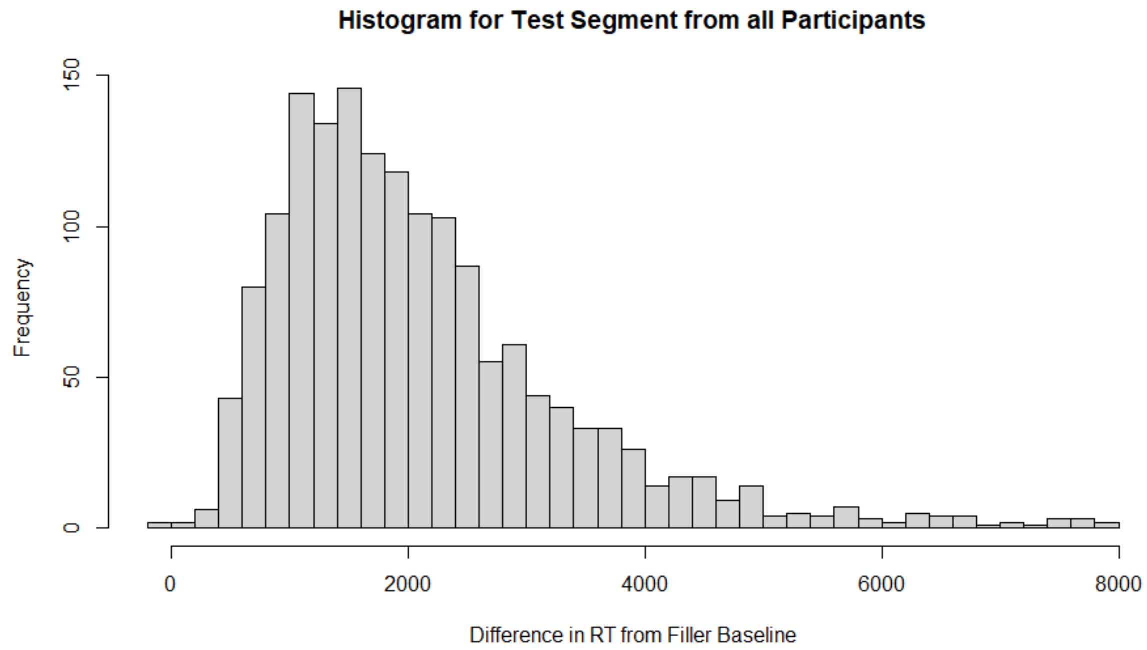

Figure 4. The histogram displays the difference in RT for the second (or test) segment in the testing sentences from the Filler Baseline sentences. Ultimately, the histogram kept 98.7% of the data.

The result is the histogram (Figure 4), from which extreme outliers have been removed. The tendency of the observation is centered around a + 1000 to +2000 difference in RTs from the Filler Baseline. Its distribution is not symmetrical, but it is still close to a normal distribution, this allowed me to use a parametric test. Its tail shows outliers getting further away from the measures of central tendency and piling up towards the right end of the range, creating a positive skew with a long tail on the right. As already mentioned, I did not want to take out all of the data points in the tail because they could be essential to the results and effects. When it comes to RT data, it is not possible to get exactly a normal distribution, because of the variation between participants.

It is worth mentioning the effects that the skewness has on the size and position of the mode, median and the mean: the greater the skewness, the greater will be the distance between the mode and the mean.

Finally, I worked with the linear mixed-effects model, the lme4 package for lmer, a two-sided linear formula object describing both the fixed-effects and random effects part of the model, (Douglas, 2015), to estimate variance for the Filler Baseline, using the following formula:

modelfill <-lmer(Diff.Filler ~Language+Order+Gender+Education+Age+RightOrLeft + (1 | Event) + (1 | Participant))

I have decided to work with the linear mixed-effects model mainly because, as opposed to a simple linear model, it allows both fixed and random effects. Moreover, the mixed effects models are useful when there is data with more than one source of random variability, like in the case of the current experiment. It is worth mentioning that a “fixed variable” is one that is assumed to be measured without error. It is also assumed that the values of a fixed variable in one study are the same as the values of the fixed variable in another study. “Random variables” are assumed to be values that are drawn from a larger population of values and thus will represent them. You can think of the values of random variables as representing a random sample of all possible values or instances of that variable (Kreft & de Leeuw, 1998).
After running this model, the summary (Table 7) showed that the Means of Intercept, Right or Left, and Order have a significant difference from the baseline. The intercepts contain every factor that is not listed underneath the intercepts itself, such as Language: English, Gender: Female, Age: 18-25 etc. The model shows that the factors of age, gender and education had no significant effects on the difference from 0, thus they were taken out from the analysis. It is worth mentioning that V2 as a factor was not included in the analysis, because of the way I have decided to split the dataset, thus it was more logical to isolate V2 effect specifically and create a dataset for it, instead of V2 being a factor itself with multiple levels. The following model is the one I used as the definitive one, and contains only significant factors such as Language , RightOr Left and Order.

modelfill <-lmer(Diff.Filler ~ Language+Order+RightOrLeft + (1 | Event) + (1 | Participant))

Through the model I would like to explain the difference in the measurement of the variables in the Filler Baseline by Language, Order of presentation and also if the correct answer was presented on the right or the left. Moreover I assume that Event (the test segment) and Participant to be sources of variance. The experiment is relatively small and it is not possible to estimate the source of variance in detail, and it does not add to the model.
The model shows that there are significant effects in Intercept (Language English and ROLLeft) with a p=<2e-16***, Order with a p=<2e-16*** and RightorLeftRight, which

choice was the correct one, with a p=0.0513·, thus it tends towards significance. Language does not have a significant effect, thus the RT is not affected by which language is the test sentence code-switched to or from. In other words, the RT is not affected by which Language condition is present in the sentence.

Min	1Q	Median	3Q	Max
-5.4	-0.6	-0.1	0.4	5.5

Table 5. This table displays the scaled residuals for the Filler Baseline to Isolate V2.

Groups	Name	Variance	SD
Event	(Intercept)	57309	239.4
Participant	(Intercept)	675375	821.8
Residual		804788	897.1

Table 6. This table displays the random effects of Event and Participant for the Filler Baseline to Isolate V2. It also shows their Variance and SD.

#Number of obs: 1610, groups: Event, 68; Participant, 24

	Estimate	Std. Error	DF	t-value	Pr(>|t|)
(Intercept)	2842.5	184.6	32.2	15.4	<2e-16 ***
LanguageNorwegian	-71.7	73.3	64.8	-1.0	0.3314
Order	-3.3	0.2	1563.6	-13.3	<2e-16 ***
RightOrLeftRight	-145.6	73.3	64.8	-2.0	0.0513·

Table 7. This table displays the fixed effects for Filler Baseline to Isolate V2.

 # Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '·' 0.1 ' ' 1**

The intercept factor shows a significant difference from the baseline when the conditions are Language English and the correct answer on the left. Order shows also significance, thus there is a learning effect. The participants do not know what they are doing at the beginning, even if the training sentences were used. They get more certain the more they go on with the experiment.

The RTs estimate compared to the baseline were obtained by adding the estimate of each factor level to the intercept to get the estimate for each overall combination of condition levels:

1. Intercept - (LanguageNor + Order + RoLLeft) = 2842.5 - 71.7 - 3.3 = 2767.5
2. Intercept - (LanguageNor + Order + RoLRight) = 842.5 - 71.7 - 3.3 - 145.6= 621.9
3. Intercept - (LanguageEng + Order + RoLLeft) = 2842.5 - 3.3 = 2839.2
4. Intercept - (LanguageEng + Order + RoLRight) = 2842.5 - 3.3 - 145.7= 2693.6

Next, I have created a boxplot for the Difference from Filler Baseline between English and Norwegian (Figure 5), in order to have a visual representation of the variability of my datasets. The black line is the median in the box plot, 50% of the data is within the gray box, while in the upper 25% of the data there are single data outliers. It is worth mentioning that the median is almost the same in both conditions, thus the boxplot can also reinforce non-significant effects of language due to lack of differences.

The English boxplot shows us that its median is closer to the bottom of the box, and that the

whisker is shorter on the lower end of the box, thus the distribution is positively skewed.
The Norwegian boxplot shows us that its median is in the middle of the box, and the whiskers are about the same on both sides of the box, thus the distribution is symmetric.
It is worth mentioning that since the interquartile ranges (that is, the box lengths) are quite small in both conditions, the data is less dispersed. Moreover, the effect of Language was not significant and it is possible to see that in the plot.

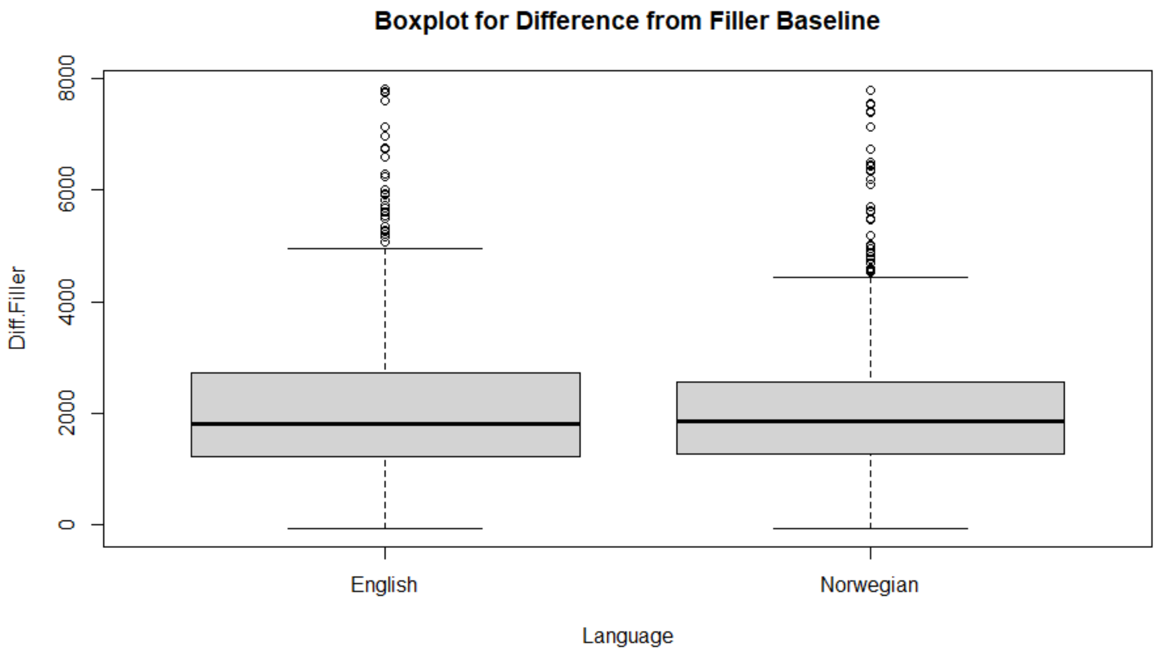

Figure 5. This boxplot is a visual representation of the variability in my datasets. It represents the Difference from Filler Baseline between English and Norwegian

4.4 Monolingual Baseline to Isolate CS

The distribution of the difference in RTs in the Monolingual Baseline Sentences is shown is the histogram (Figure 6), and it was created with the following Formula:

histmono <-hist(MonoBL$Diff.Mono, main="Histogram for Test Segment from all Participants", xlab="Difference in RT from Monolingual Baseline", breaks = 50)

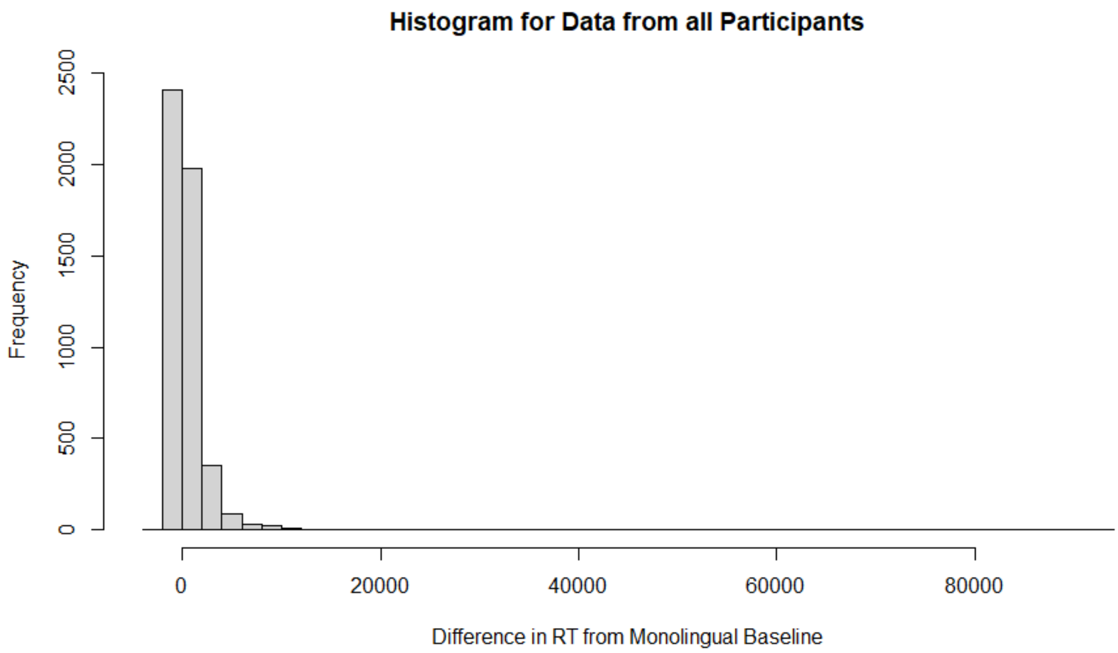

Figure 6. This histogram displays the distribution of the difference in RTs in the Monolingual Baseline Sentences.

From the summary of the histogram in RStudio, I noticed that the values in the bins (columns in the histogram), had three very high data points, making the x axis stretch out. There seems to be other outliers, making a long right tail. In order to obtain a less stretched out histogram, and since I was interested only in the test segment of the filler sentences, I have picked out only the second segment of the filler sentences (Figure 7).

I removed outlier points for “Difference” and leveled out its ends by firstly calculating the “Mean” of the “Difference” column in the dataset. Then, I calculated the two “Standard Deviation” of the values in the “Difference” Column, which is “5580.566”. Finally, I added the two “Standard Deviation” to the “Mean” (the mean is 522.0) = 6102.5, to obtain the upper boundary. I also removed the 2 SD from the “Mean” (-5058.6) in order to get the lower boundary. As already mentioned in the previous paragraph, I have decided to keep all “Difference” values that are within two SD from the mean. The result is a Monolingual Baseline Histogram (Figure 7), only for the test position, using the following formula:

histmonoseg2 <-hist(MonoBLseg2$Diff.Mono, main="Histogram for Test Segment from all Participants", xlab="Difference in RT from Monolingual Baseline", breaks = 50)

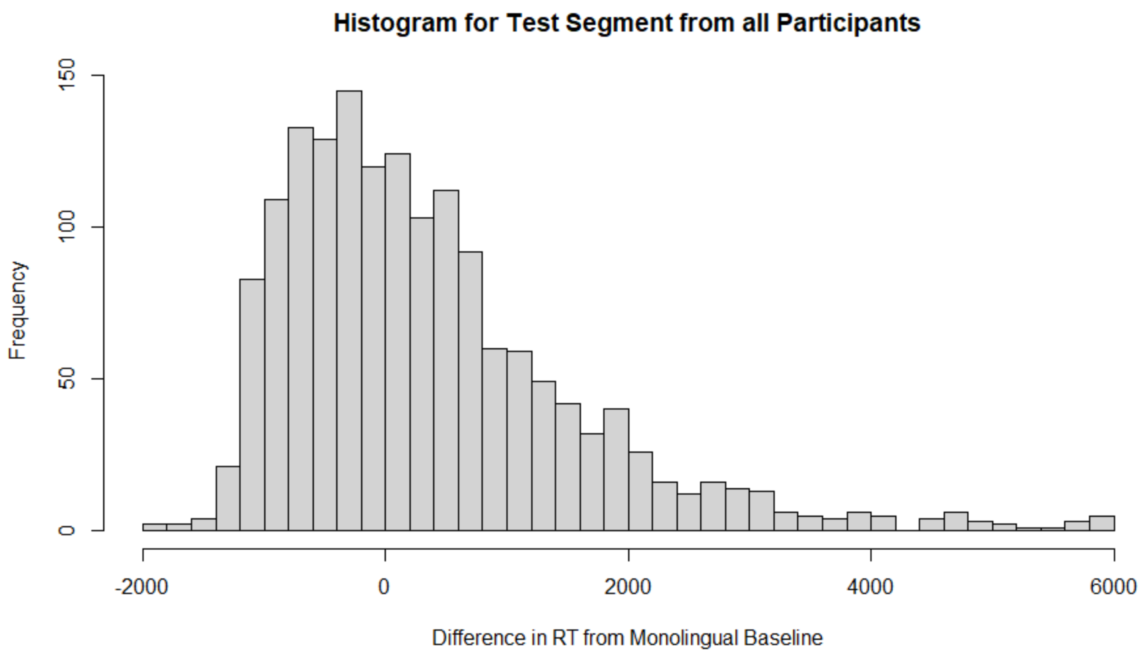

Figure 7. The histogram displays the difference in RT for the second (or test) segment in the testing sentences from the Monolingual Baseline. Ultimately, the histogram kept 98.6% of the data.

The result is an histogram (Figure 7), from which extreme outliers have been removed. The tendency of the observation is centered around a -1000 to +1000 difference in RTs from the Monolingual Baseline. Similarly to the other baseline, the distribution is not symmetrical, but it is still close to a normal distribution, which allowed me to use a parametric test. Its tail shows outliers getting further away from the measures of central tendency and piling up

towards the right end of the range, creating a positive skew with a long tail on the right.
Finally, I have worked with the linear mixed-effects model to estimate variance for the Monolingual Baseline, using the following formula:

modelmono <-lmer(Diff.Mono ~ Language+Gender+Education+Age+RightOrLeft + (1 | Event) + (1 | Participant))

Similarly to its counterpart with the filler sentences as baseline, this model also does not show that the factors of age, gender and education have had significant effects on the difference in RTs, so I did not include them in the analysis.
The following model is the one I used as the definitive one, and contains only significant factors such as Language , RightOr Left and Order:

modelmono <-lmer(Diff.Mono ~ Language+Order+RightOrLeft + (1 | Event) + (1 | Participant))

The model's summary (Table 10) shows that there are significant effects in Intercept (p=4.64e-06***), Order (p= < 2e-16***) and RightOrLeft (p=0.05·), as follows:

Min	1Q	Median	3Q	Max
-5.4	-0.6	-0.1	0.4	5.5

Table 8. This table displays the scaled residuals for Monolingual Baseline to Isolate CS.

Groups	Name	Variance	SD
Event	(Intercept)	58243	241.3
Participant	(Intercept)	662626	814.0
Residual		798864	893.8

Table 9. This table displays the random effects of Event and Participant for the Monolingual Baseline to Isolate CS. It also shows their Variance and SD.

#Number of obs: 1609, groups: Event, 68; Participant, 24

	Estimate	Std. Error	DF	t-value	Pr(>|t|)
(Intercept)	1005.1	183.2	32.4	5.5	4.64e-06 ***
LanguageNor	71.4	73.6	64.8	0.9	0.3355
Order	-3.2	0.2	1561.9	-13.3	<2e-16 ***
RightOrLeftRight	141.7	73.6	64.8	-1.9	0.0585·

Table 10. This table displays the fixed effects for the Monolingual Baseline to Isolate CS.

#Signif. codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '·' 0.1 ' ' 1**

The intercept factor shows a significant difference from the baseline when the conditions are Language English and the correct answer on the left. Order shows also significance, thus there is a learning effect. The RTs estimate compared to the baseline are:

1. Intercept - (LanguageNor + Order + RoLLeft) = 1005.1 + 71.4 - 3.2 = 1073.3
2. Intercept - (LanguageNor + Order + RoLRight) = 1005.1 + 71.4 - 3.2 + 141.7 = 1214.9
3. Intercept - (LanguageEng + Order + RoLLeft) = 1005.1 - 3.2 = 1001.9
4. Intercept - (LanguageEng + Order + RoLRight) =1005.1 - 3.2 + 141.7 = 1143.5

I have then created a boxplot for the Difference from Monolingual Baseline between English and Norwegian (Figure 8). The black line is the median in the box plot, 50% of the data is within the gray box, while in the upper 25% of the data there are single data outliers. As for the Filler Baseline, the median is almost the same in both conditions, and the boxplot can also reinforce non-significant effects of language due to lack of differences.

Similar to the analysis of the Filler Baseline, in the Monolingual Baseline, the English boxplot shows us that its median is closer to the bottom of the box, and that the whisker is shorter on the lower end of the box, thus the distribution is positively skewed.

The Norwegian boxplot shows us that its median is in the middle of the box, and the whiskers are about the same on both sides of the box, thus the distribution is symmetric. Also in this case, the interquartile ranges are quite small in both conditions, thus the data is less dispersed.

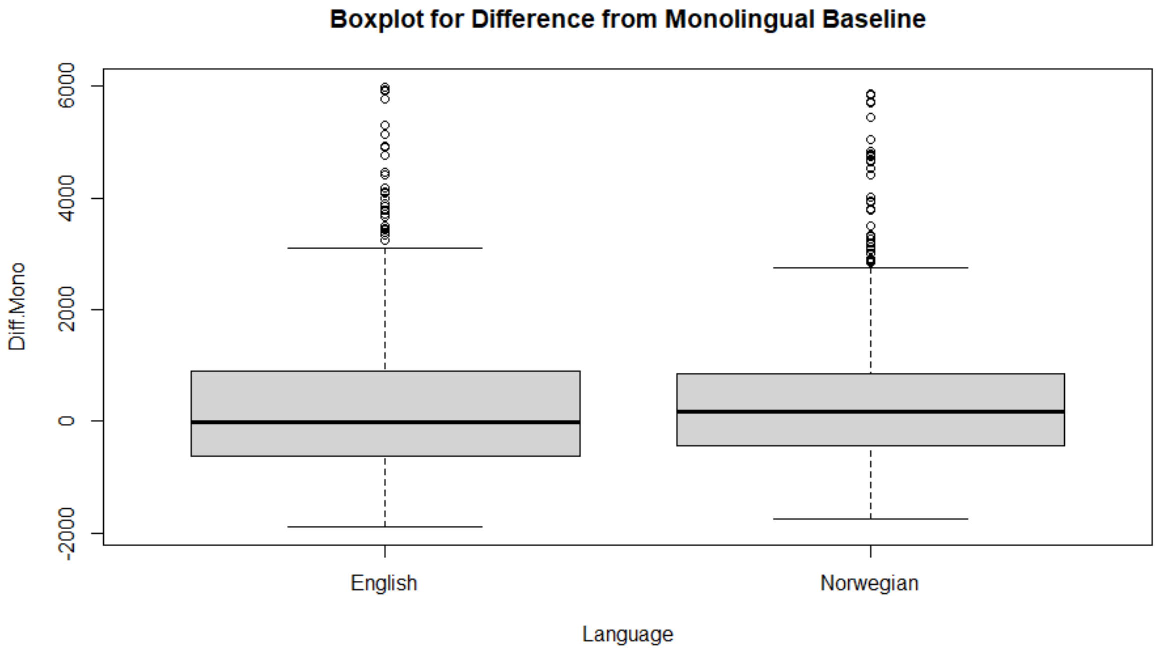

Figure 8. This boxplot is a visual representation of the variability in my datasets. It represents the Difference from Monolingual Baseline between English and Norwegian

4.5 Learning Effect Due to Order of Presentation for Both Datasets

Before mentioning anything else, it is important to state that originally the learning effect was not included in the model (but order was included as a factor). However, I decided to regress it out to get clearer results for the effects.

I have decided to take a closer look at the order factor and the learning effect through ggplots, since it is possible to notice that the estimated effect of the slope of the order looks consistent and similar in both models, regardless of the baseline. It is possible that it stays consistent, but I will delve into it in the discussion section.

On average the participants got about 3 milliseconds faster for each exposure — this is highly significant (i.e. cannot be explained by uncertainty in the data). The RightOrLeft effect got smaller, as well as Language, and according to the Order factor (which takes a lot away from RightOrLeft), there was a learning effect, which is independent and increases with training. The learning effect would not go on endlessly, but as already mentioned, I have accounted for it, by including it as a factor in the model (Order). As shown in the following plots, I isolated the learning effect in both the monolingual baseline (Figure 9) and filler baseline (Figure 10) datasets. I have used the ggplot2 package and the codes (17,18) to plot a fitted linear regression model in order to visualize the learning effect due to order of presentation respectively for the monolingual baseline (17) and the filler baseline (18), by using the following R syntax:

(17)

```
ggplot(data=MonoBLseg2, aes(x = Order, y = Diff.Mono, color = Language)) +
geom_smooth(method = "lm") +
labs(title = "Learning Effect Due to Order of Presentation (Monolingual Baseline)",
x = "Order",
y = "Difference in RT")
```

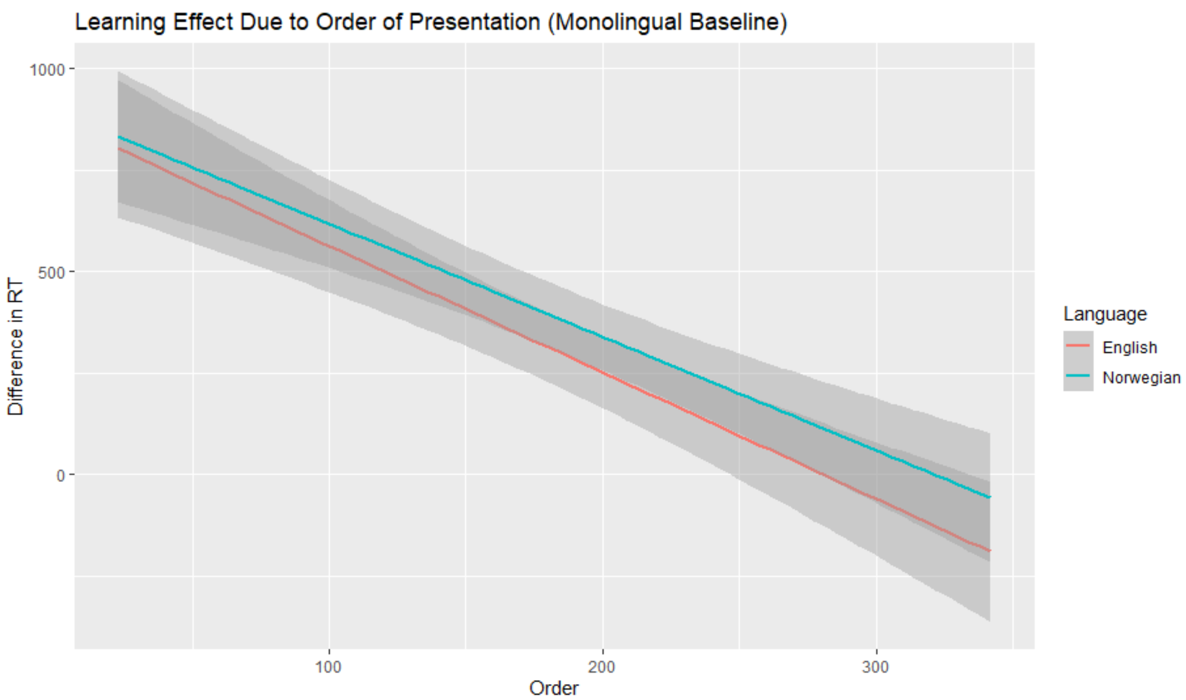

Figure 9. This table shows the learning effect due to order of presentation for the monolingual baseline.The x-axis represents the Order, while the y-axis represents the Difference in RT. The slope language per order looks really the same and within the confidence intervals, shown by the gray area.

(18)

```
ggplot(data=FillerBLseg2, aes(x = Order, y = Diff.Filler, color = Language)) +
geom_smooth(method = "lm") +
labs(title = "Learning Effect Due to Order of Presentation (Filler Baseline)",
x = "Order",
y = "Difference in RT")
```

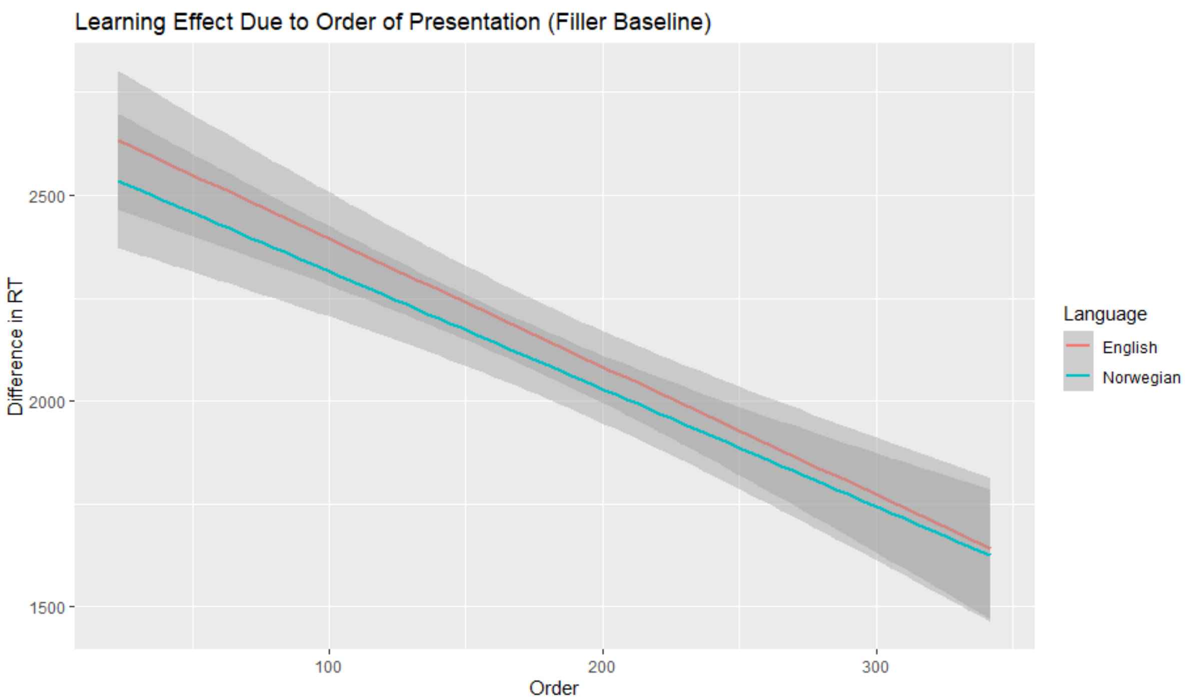

Figure 10. This table shows the learning effect due to the order of presentation for the filler baseline. The x-axis represents the Order, while the y-axis represents the Difference in RT. The slope language per order looks really the same and within the confidence intervals, shown by the gray area.

These simple regression models (17 and 18) show that the learning effect is similar in both language conditions (Figure 9 and Figure 10), and that the line has a negative slope, so it is showing a learning effect where the participant gets faster over time after being exposed to a greater number of presentations (order). The difference in the y-axis is present because the testing sentence RT differs overall by a greater amount from 0 with the filler baseline than with the monolingual one. I am still using the difference in RT on the y-axis, so it will be different from 0 since there is an effect of V2 (filler baseline) or CS (monolingual baseline).

5. Discussion

5.1 Discussion of the Research Questions and Hypotheses

Based on the results in the previous chapter I will take a look at my hypothesis and evaluate whether or not they are supported by my data. For the convenience of the reader, I present and discuss the research questions and hypotheses in this section as well.

5.1.1 The Null Hypothesis

The null hypothesis is that there is no variation between the sentence conditions.
I found an effect of CS regardless of the language being switched to. In addition, a significant difference from baseline for V2 was also detected. Thus, is it possible to reject the null hypothesis? The answer would be both yes and no. There was no effect based on the factor of Language, but there is definitely a longer RT compared to the baseline.
The first research question seeks a difference in RT when processing the code switched sentences compared to a baseline without the CS. According to the data, a difference in RT was detected, but as already mentioned, the variation from the baseline exists regardless of the language condition.

5.1.2 The First Alternative Hypothesis

My first alternative hypothesis is that the code-switching demands greater processing, thus a longer reaction time. The RTs will be faster when the participant will switch from L2 English to L1 Norwegian. This is due to the fact that even if their L2 is strong, the L1 is still the dominant language system, and it should be an advantage to switch to it.
It is not possible to fully reject this hypothesis based on the data I have now, since the RT is significantly different from the monolingual baseline. Therefore, it needs to be tested in the future.

I found evidence that the CS increases the RT, but the data does not support a significant difference based on the factor of Language. The previous study conducted by Schwochow, Larsen and Johansson (2021) predicted faster RT when switching to L1, thus it needs further experimenting.

5.1.3 The Second Alternative Hypothesis

My second alternative hypothesis is that the V2 word order demands greater processing due to a grammar shift. Therefore, the sentences that contain both CS and V2 will be even more difficult to process, resulting in a higher RT. If CS needs to activate a second grammar, this can be predicted to make a reparse involving cross-linguistic input more difficult and thus adding CS in a V2 would lead to slower reaction times.

To answer the second research question: in which sentence condition is there a difference in RT to the baseline when the participant is faced with the choice at the testing position in a code-switching environment that also meets the V2 requirement in L1 Norwegian? My data support that the V2 effect has an estimate for the intercept of 2800ms, while the monolingual baseline used to isolate the CS effect only had 1000ms. Thus, both V2 and CS add more processing time (see Table 7 and 10).

5.1.4 The Third Alternative Hypothesis

My third and last alternative hypothesis, additionally investigates the learning effect due to the presentation order of the sentences in the experiment. The implications of the learning effect is that the participants get faster towards the end of the experiment, and it is not related to the sentence conditions. As possible to see from Figure 9 and 10, there is a small and significant learning effect in both filler and monolingual baseline, that adds over time.

5.2 Filler Baseline Results V2

I found that V2 takes longer to process compared to the baseline without the V2, according to my data. There was a significant effect, regardless of the language condition of the sentence. In the previous experiment (Schwochow et al., 2021) the results suggested a faster processing time when switching to a L1 that contained the V2 rule, but the V2 was not isolated. Such results are not supported by my data. Thus, it might be possible that V2 could actually provide a disadvantage.

Moreover, the fact that it took longer to process sentences with a V2 word order could mean that both grammars are not activated at the same time, hence there is a presence of a slow down effect. When the sentence begins in Norwegian the grammar with the V2 requirement should be active. Yet, the swap to English still causes a slow down when V2 is an option, maybe because the English grammar needs to be activated. Similarly, when the sentence starts in English, the grammar with no V2 requirement should be active. However, the swap to Norwegian causes a slow down, probably because the Norwegian grammar needs to be activated.

I would like to add that Norwegian also has a SVO word order, so it might be that the V2 requirement would increase the RT because it incorporates both rules.

5.3 Monolingual Baseline Results CS

I predicted based on the results of past research that the RTs will be faster when the participant will switch from L2 English to L1 Norwegian, but this hypothesis was not supported by my data. Perhaps the advantage of switching to L1 was negated by having to fulfill the V2 requirement at that point. It is L1 grammar, but it is a big shift from L2 right at that point. It might also be due to the fact that usually most experiments force unnatural circumstances on the participant by code-switching at more anomalous points in a sentence. In real life, people do not naturally codeswitch at such points, probably it is easier.

I would like to refer back to section 1.1.5.2 and 1.1.5.3, the previous research done respectively on code-switching to a L2 and code-switching to a L1. My experiment, contrary to the just mentioned ones, is in both language directions. Moreover, there was no speed up observed for CS from L2 to L1, unlike Schwochow's experiment (et al., 2021). This might depend on the power of the experiment, or because the V2 counteracts the CS effect. In other words, the V2 transformation counteracts the switch to L1 and takes away the advantage of the L1. In my study, the V2 is canceled out by the monolingual baseline but always present in the two possible answers, thus is it possible that V2 cancels out the advantage of the L1, probably affecting the results.

Both in the filler and the monolingual baseline, the intercept (Table 7 and Table 10) shows a significant difference from the baseline, which means that CS causes an increase in RT. But there were no specific effects of which language it was code-switched to and from, only the general CS and V2 effects.

5.4 Order and Learning Effect

The Order factor takes into account the learning effect, and it stands for the order and number of presentations, including questions regarding age, gender and education, as well as training, filler and testing sentences. The participants were presented with the same sentence stimuli multiple times in different conditions. Because of it I would expect there to be a learning effect. In order to regress the learning effect out of the data, I have used a newer technique of including order as a factor in my model. The order of the presentation is always recorded for each participant even though it is randomized. If Order was not included as a factor, it would have been included in context with the effects of the other factors. On average the participants got about 3 milliseconds faster for each exposure this is highly significant (i.e. cannot be explained by uncertainty in the data). The RightOrLeft effect got smaller, as well as Language, and according to the Order factor (which takes a lot away from RightOrLeft), there was a learning effect, which is independent and increases with training. It is possible to notice that the learning effect is similar in both language conditions, the participants get faster over time after being exposed to a greater number of presentations (order). It is worth mentioning that the learning effect stays consistent for all conditions, since the estimated effects for each model are -3.2ms for the filler baseline (V2) and - 3.3ms for the monolingual baseline (CS). This tendency has been detected also in other experiments conducted in the same laboratory, such as the research of Larsen and Johansson (2022), and Salte (2022), which had similar learning effect size. This could mean that learning effects can be regressed out of other experiments, particularly RT ones, which lead to clearer effects.

5.5 Right or Left

Another significant difference was detected for the factor of Right or Left. First of all, the reason why I have decided to account for it was to have a balanced presentation of the stimuli. From my results, it is possible to notice that the presentation of the correct option on the right is faster in the filler baseline. The reason might be that in order to select the choice on the right, the participants need to read only left and then right. Yet to select the left choice, they might read the left choice first, then the right one and left again.
The RightorLeftLeft is faster (141.7ms) in the monolingual baseline, probably because they select it if correct, without looking at both options. While it is slower (-145.6ms) in the filler baseline, probably due to the fact that the participants need to check both answers because the V2 takes more computational effort, thus they need to do more back and forth.

6. Conclusion

This study has investigated the question on whether or not there is a difference in RT to the baseline when the participant is faced with the choice at the testing position in a code-switching environment that also meets the V2 requirement in L1 Norwegian? Therefore, I explored the effect of CS to L1 Norwegian and L2 English in V2 word order sentences at the point where the V2 requirement needs to be implemented in Norwegian, but not in English. My data supports that both CS and V2 had longer reading times compared to a baseline without CS or V2 on the individual conditions, regardless of the language condition.

In short, CS did add extra difficulty for reading or deciding on the testing sentences, but the data does not support a significant difference based on the factor of Language. A significant increase in RT was also detected when switching to L1 or L2 due to V2.

A possible explanation to consider in future research is the fact that the V2 transformation counteracts the code switch to L1. I would like to duplicate the analysis in the future with a greater number of filler sentences to use as the baseline, by controlling them more and by using different structures that do not have the V2 confusion.

As already mentioned, in order to regress the learning effect out of the data, I have used a newer technique of including order as a factor in my model. The order of the presentation is always recorded for each participant even though it is randomized. Future research should try to incorporate this method, also in order to follow the decision making process in more detail. Therefore, a good design choice would be to include the presentation order per participant in future studies to see if it affects the results, since it has been found to have a significant effect on the RT.

7. Bibliography

Berger, A., & Kiefer, M. 2021. Comparison of Different Response Time Outlier Exclusion Methods: A Simulation Study. Frontiers. Section for Cognitive Electrophysiology, Department of Psychiatry, Ulm University, Ulm, Germany. DOI: https://doi.org/10.3389/fpsyg.2021.675558

Bryman, A., & Bell, E. 2011. Business Research Methods. Oxford: Oxford University Press. Retrieved from https://research-methodology.net/research-methodology/ethical-considerations/#_ftn1.

Bullock, B. E. & Toribio, A. J. (Eds.). 2009. The Cambridge handbook of linguistic code-switching. Cambridge University Press. DOI: 10.1017/CBO9780511576331

Creswell, J. 2009. Research Design: Qualitative, Quantitative, and Mixed-Method Approaches, Sage Publications. Retrieved from https://www.researchgate.net/publication/225083951_Research_Design_Qualitative_Quantitative_and_Mixed-Method_Approaches

Dillon, B. 2009. Grammatical illusions in sentence processing: At the interface of performance and competence, University of Massachusetts, Amherst. Retrieved from https://www.ling.upenn.edu/Events/PLC/plc40/abstracts/DillonPLC40Keynote.pdf address.

Fischer, O., Van Kemenade, A., Koopman, W. & Wurff, W. 2001. The Syntax of Early English. DOI: 10.1017/CBO9780511612312.

Forster, K. I. 2010. Using a Maze Task to Track Lexical and Sentence Processing. The Mental Lexicon Methodological and Analytic Frontiers in Lexical Research (Part I) 5 (3): 347–57. DOI: 10.1075/ml.5.3.06wur.

Heller, K., Matsak, E., Abboud, H., Schultz, H., & Zeitlin, V. 2015. Superlab 5 (version 5.05). English. Cedrus corporation.

Johansson, C., Rugaard, E. & Asperheim, H. 2019. Low cost of Code Switching. DOI10.13140/RG.2.2.29763.35362.

Kreft, I., & de Leeuw, J. 1998. Introducing multilevel modeling. Sage Publications, Inc. https://doi.org/10.4135/9781849209366

Nicol, J., & Swinney, D. 1989. The role of structure in coreference during sentence comprehension. Journal of psycholinguistic research. 18. 5-19. 10.1007/BF01069043.

Odlin, T. 1989. Language Transfer: Cross-Linguistic Influence in Language Learning. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139524537.

Putnam, M., Carlson, M., & Reitter, D. 2018. Integrated, not isolated: Defining typological proximity in an integrated multilingual architecture. Frontiers in Psychology, 8, 2212. DOI: 10.3389/fpsyg.2017.02212.

R Core Team. 2021. R version 4.1.2 (2021-11-01). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Salte, A. 2022. Stay way away. Grammaticalization of way. MA thesis, University of Bergen.

Saunders, M., Lewis, P., & Thornhill, A. 2012. “Research Methods for Business Students” 6th edition, Pearson Education Limited.

Schwochow, S., Larsen, T., & Johansson, C. 2021. The bilingual parser predicts one structure: code-switching in the garden path. DOI: 10.13140/RG.2.2.20940.95368.

Sportiche, D., Koopman, H., Stabler, E. 2014. An Introduction to Syntactic Analysis and Theory. Wiley Blackwell. pp. 68–70, 189–191. ISBN 978-1-4051-0017-5.

Larsen, T. & Johansson, C. 2022. Control in a Norwegian Grammar Maze. Manuscript submitted for publication.

Townsend, D. J. & Bever, T. G. 2001. "Embedding the Grammar in a Comprehension Model". Sentence Comprehension: The Integration of Habits and Rules. Cambridge, MA: MIT Press. pp. 157–185.

Urk, C. 2020. "Verb Second is syntactic", Rethinking Verb Second, Oxford University Press, pp. 623–641, doi:10.1093/oso/9780198844303.003.0026, ISBN 978-0-19-884430-3.

Wager, M., Lau, E. F. & Philips, C. 2009. Agreement attraction in comprehension: Representations and processes. Retrieved from https://www.researchgate.net/publication/222432652_Agreement_attraction_in_comprehension_Representations_and_processes address.

Warren, Paul. 2013. Introducing Psycholinguistics, Cambridge Introductions to Language and Linguistics. p. 4-5.

Westergaard, M. 2003. Unlearning V2: Transfer, markedness, and the importance of input cues in the acquisition of word order in English by Norwegian children. *Eurosla Yearbook, 3*, 77-101. DOI: 10.1075/eurosla.3.07wes

Zentella, A. C. 1997. Growing Up Bilingual: Puerto Rican Children in New York. Malden: Blackwell Publishers. DOI: 10.1075/lplp.21.3.23mar

8. List of Appendices

Appendix A - Dataset

Training sentences:

1. Norge har en lang felles grense med Sverige.
   Norway has a long common border with Sweden.
2. Ingen så at det var maling flekker.
   No one saw that there were paint stains.
3. Hunden min var hjemme da jeg våknet i dag.
   My dog was home when I woke up today.
4. Katten hoppet ut av vinduet og opp på taket i går.
   The cat jumped out the window and onto the roof yesterday.
5. Restauranten aksepterer ikke American Express -kort.
   The restaurant does not accept American Express cards.
6. Mange butikker i Bryggen selger suvenirer og kart.
   Many shops in Bryggen sell souvenirs and maps.

Testing sentences:

1. Etter jobb spiser jeg middag på restauranten med mine venner.
   After work I eat dinner at the restaurant with my friends.
2. Når jeg kommer hjem på fredager bestiller vi pizza og ser på en film..
   When I get home on Fridays, we order pizza and watch a movie.
3. Da jeg var barn, bodde jeg i Oslo sammen med besteforeldrene mine.
   When I was a child, I lived in Oslo with my grandparents.
4. Hvis du gir bilen til henne, føler jeg at det ikke er rett.
   If you give the car to her, I feel that it is not right.
5. Før klokka åtte spiser han pannekaker og bacon til frokost.
   Before eight o'clock he eats pancakes and bacon for breakfast.
6. Så lenge været er fint, må jeg ta hunden ut på tur.
   As long as the weather is nice, I must take the dog out for a walk.
7. Mens jeg lagde middag til hele familien, dekket hun bordet.
   While I made dinner for the whole family, she set the table.
8. Fordi han føler seg dårlig og utlistt, holder han seg hjemme.
   Because he feels bad and exhausted, he stays at home.

9. Siden det var så fint vær, spiste vi i hagen med naboene.
   Since the weather was so nice, we ate in the garden with the neighbors.

10. Ettersom du lytter til musikk på Spotify, får du bedre anbefalinger.
    Since you listen to music on Spotify, you get better recommendations.

11. Dersom elevene får god selvtillit, blir de flinkere i fagene.
    If the students gain good self-confidence, they become better at the subjects.

12. De dagene det er fint vær, rir jeg på hesten i stallen.
    On days when the weather is nice I ride the horse at the stables.

13. Selv om det var rødt lys, kjørte han rett gjennom.
    Even though there was a red light, he drove straight through.

14. Til tross for at de alle var nederlendere, snakket de tysk med hverandre.
    Despite the fact that they were all Dutch, they spoke German to each other.

15. Enda gutten ikke kunne svømme, hoppet han ut i elva.
    Even though the boy could not swim, he jumped into the river.

16. Hvis du har lyst til å gå på kino, foreslår jeg at vi ser en musikal.
    If you want to go to the movies, I suggest we watch a musical.

17. I min barndom bodde jeg i Bergen i omtrent 10 år.
    During my childhood, I lived in Bergen for around 10 years.

Filler sentences:

1. Datamaskinen min ville ikke kjøre programmet denne morgenen.
   My computer would not run the program this morning.
2. Professoren introduserte et nytt tema for studentene.
   The professor introduced a new topic for the students.
3. Anna kom ikke til timen i dag, fordi hun var syk.
   Anna did not come to the appointment today because she was ill.
4. Mannen som var sur spilte tennis med sin venn.
   The man who was angry, played tennis with his friend.
5. Torbjørn gjørde ikke hjemmeleksa si denne uken.
   Torbjørn did not do his homework this week.
6. Lars var bedre i sjakk enn noen andre i byen.
   Lars was better at chess than anyone else in the city.
7. Treet vokste seg større enn huset til naboen.
   The tree grew larger than the neighbor’s house.

8. Maleriet på veggen til Marie avbildet masse folk som danset.
   The painting on Marie's wall depicted many people dancing

Appendix B - Digital Reading and Decision Task

Digital Reading and Decision Task

Participant Information Sheet

This is an inquiry about participation in a research project.

Who is responsible for the research project?

The HumLab at the University of Bergen is the institution responsible for the project. Annalisa Arcidiacono is the individual researcher responsible for the planning and execution of the project.

What does participation involve for you?

Participation in this study will involve one session.

The study consists of a simple decision task. The responses will anonymously be recorded and statistically analyzed. This session will take approximately 10-15 minutes.

Disadvantages/risks of participation

Your participation in this study does not involve any physical or emotional risk to you beyond that of everyday life.

Participation is voluntary

Participation in the study is voluntary. If you chose not to participate, you can withdraw your consent at any time without giving a reason. There will be no negative consequences for you if you choose not to participate.

Your personal privacy – how we will store and use your personal data

Gender and age information will be collected; however, no personal data enabling participant identification will be collected for this study. We will use the reaction time data obtained from your participation solely for the purpose of this particular study.

The results of this study will be used in publications and presentations, including a master thesis and possible PhD dissertation.

Compensation

You will not be compensated for your participation.

Where can you find out more?

If you have questions about the project, please contact:

- Annalisa Arcidiacono via Annalisa.Arcidiacono@student.uib.no
- Tori Larsen via Tori.Larsen@uib.no
- Prof. Christer Johansson via Christer.Johansson@uib.no

Yours sincerely,

Project Leader

Consent form

I have read and understood the information about the digital reading and decision task.
I understand that my participation is voluntary and that I am free to withdraw at
any time without having to give a reason and without facing any negative
consequences.
I confirm that I am at least 18 years old.
I confirm that I am a university student.
I give consent to participate in this project and for my reaction time data to be collected for analysis.

Name of participant Date Signature

Appendix C - Experiment's Instructions

Du må ta en avgjørelse mellom to fortsettelser. Velg den som føles best for deg.

Bruk venstre (RØD) eller høyre (GRØNN) knapp for å bestemme mellom venstre og høyre setning. Du må ta en avgjørelse på delene med brakettene.

Trykk på knappen i midten formet som en sirkel når du ikke trenger å ta en beslutning om den delen av setningen.
Eksempel:

Det er som   å stjele godteri fra et   [barn ||vaskeklut]

Vi ville ha trykket på knappen i midten for de to første delene av setningen og vi ville ha trykket på venstre knapp for «barn».

Når du har startet eksperimentet du gir samtykke til å delta i denne studien.

All deltakelse er frivillig og anonym. Du kan forlate rommet når som helst i løpet av eksperimentet hvis du vil avbryte eksperimentet.

Trykk på hvilken som helst knapp for å starte eksperimentet og for å bekrefte at du er en universitetsstudent som er norsk som morsmål.

Hva er ditt kjønn?
kvinne - trykk på venstre knapp
mann - trykk på høyre knapp
annet - trykk på den øverste knappen
foretrekker å ikke svare - trykk på den nederste knappen

Hvor gammel er du?
18-25 - trykk på venstre knapp
26-35 - trykk på høyre knapp
36-45 - trykk på den øverste knappen
46+ - trykk på den nedeste knappen
foretrekker å ikke svare - trykk på den midterste knappen

Hva er utdanningsnivået ditt?

videregående skole - trykk på venstre knapp

bachelorgrad - trykk på høyre knapp

mastergrad - trykk på den øverste knappen

annet - trykk på den nedeste knappen

foretrekker å ikke svare - trykk på den midterste knappen