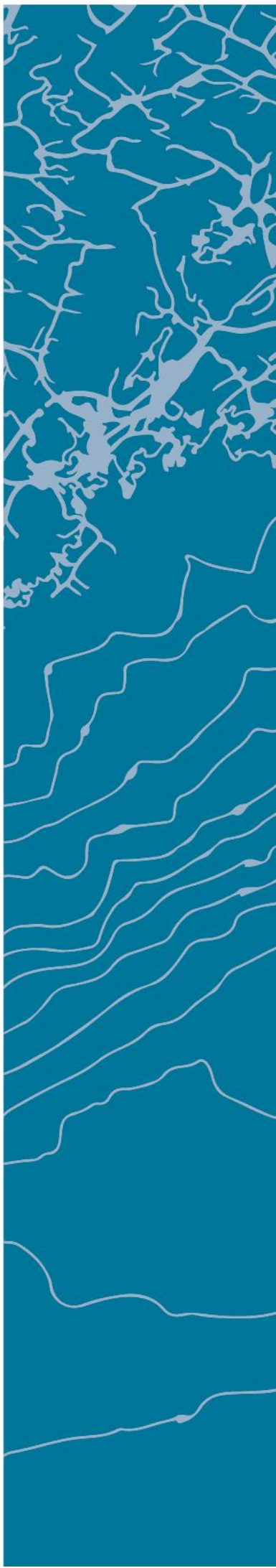
UNIVERSITETET I AGDER

Word Neighbor Priming in Bilingual Sentence Reading: Evidence from Eye Movements

DAG MAGNE HAUGLAND

SUPERVISOR
Linda Ruth Wheeldon

University of Agder, 2018
Faculty of Humanities and Education
Department of Foreign Languages and Translation



Contents

Abstract 1
Introduction 2
Effects of form similarity in first language word reading 3
Form similarity in sentence reading: Effects of proficiency 4
Reading processes in bilinguals 6
Models of non-selective language activation and suppression 6
Evidence for the BIA model 8
Evidence for the RHM model 10
The Inhibitory Control Model 12
L2 sentence reading and eye-tracking 13
Bilingual profile and L2 proficiency 13
Language comparison 17
Phonetic inventory 18
Phonetic inventory comparison 20
Inflection 20
This study – experiments, motivations, questions and predictions 21
Key manipulations 22
Method 23
Materials 23
Design 24
Apparatus 25
Procedure 25
Participants 25
Results: Eye-tracking experiment 26
Short conditions 26
First fixation 26
Regression-path 27
Total time 27
Long conditions 27
Total time 28
Discussion: Eye-tracking experiment 28
Results: Correlations with Principal Components of bilingual profile 30
Correlations for the short conditions 31

Correlations for the long conditions 31
Discussion: Correlations with PCA factors 32
General discussion 33
Conclusion 35
References 36
APPENDIX 39

Abstract

This study investigates English sentence reading processes in Norwegian-English bilinguals. It aims to explore the effects of word form overlap and its connection to second-language (L2) English proficiency. We recorded eye movements of Norwegian-English bilinguals while they read single sentences on screen. The experiment was a replication of Frisson, Koole, Hughes, Olson and Wheeldon, (2014), who investigated the effects of form similarity between words (orthographic and phonological, both separate and combined) on sentence reading in English monolinguals. They also investigated the effects of distance between prime and target words, and found that words that overlap in both orthography and phonology cause inhibition (longer gaze durations) when the distance between prime and target was 3 words. Interestingly only skilled readers showed the same effect when distance between prime and target was 9 words. Our study aimed to investigate whether the same inhibitory effects occur in Norwegian-English bilinguals and to determine if aspects of L2 proficiency modulated the effect.

In our experiment, the sentences contained a prime word (related) or matched control word (unrelated) and a target word. In the related conditions, the prime and target overlapped both orthographically and phonetically in the end (e.g. *fork-pork*) or the beginning (e.g. *rail-raid*). We also examined the effects of distance between prime and target. In order to investigate effects of individual differences in bilingual language profile and proficiency, the results were then correlated with factors of a bilingual profile established using data from a revised LEAP-Questionnaire from a sister project.

We found that proficient L2 users show more inhibition when there is phonological and orthographic overlap. Generally, there is inhibition in end overlap items, and facilitation in begin overlap items, but only when the distance is about 3 words. When prime and target is separated by 9 words, the effect disappears. We also found that there is a significant correlation between L2 proficiency and the effects of overlap. The more proficient the bilingual is in their L2, the more native-like the inhibition effects. We conclude that proficient L2 readers are more likely to be affected by word overlap priming, and that skilled L2 readers are negatively impacted by word overlap to a higher degree than less skilled L2 readers are. These findings suggest that, similar to L1 readers, skilled L2 readers maintain word-form information for longer when reading than less skilled readers.

Introduction

This study investigates the effects of word form similarity (orthographic and phonological) on the reading of second language English by native Norwegian speakers. The processes of word reading in a native language (L1) have been researched extensively. In single word reading paradigms, there is evidence of interference from orthographically similar words (e.g. corn-horn), suggesting competition between related words during word recognition (e.g. Davis & Lupker, 2006; Forster & Davis, 1984). Sentence reading processes have also been extensively studied using eye-tracking technology (e.g. Andrews & Lo, 2012; Frisson, Koole, Hughes, Olson, & Wheeldon, 2014).

Interestingly, interference effects (e.g. longer gaze durations) have also been demonstrated between orthographically and phonologically related words occurring in the same sentence, e.g. (*"There was a blur as the blue lights of the police car whizzed down the street"* (Frisson et al, 2014). All these effects were stronger (occurred over a greater number of intervening words) for more proficient readers. These results suggest that readers that are more proficient retain low-level information for longer than less proficient readers do.

The focus of this study is on second language (L2) reading processes. Bilinguals vary in their L2 proficiency and L2 language experience (e.g. age of acquisition) and their use of the second language (Marian, Blumenfeld, & Kaushanskaya, 2007). Little is known about how such factors relate to L2 reading processes. This study therefore investigates the effect of second language (L2) proficiency and usage on L2 sentence reading processes in bilinguals.

Self-rated language proficiency has been shown to predict objectively tested language skills in bilinguals (e.g. Marian et al, 2007). The aim of this study was to determine if L2 readers show similar effects of word-form overlap to that of L1 readers, and if so, how the effects observed relate to L2 proficiency and language experience. The study involves a replication of Frisson et al (2014) with Norwegian-English bilinguals.

Additionally, detailed questionnaire data was collected in a sister study (Handeland & Zitong, 2018), through a revised edition of the Language Experience And Proficiency Questionnaire (LEAP-Q). The data collected were analyzed for this study as well in order to give a clear picture of the different factors that might affect L2 reading performance.

Before describing the methodology and results, I will first review the effects of form similarity in word reading processes in monolinguals, (e.g. Bowers, Davis, & Hanley, 2005, Davis & Lupker, 2006; Forster & Davis, 1984) and the theoretical explanations of these effects, in particular the IAM model (McClelland and Rumelhart, 1981), which is a model of lexical competition in word processing. I will also examine the effects of proficiency on single word processing (e.g. Andrews & Lo, 2012). The following section will review reading processes in bilinguals and models of non-selective language activation and suppression (e.g. Green, 1998; Sunderman & Kroll, 2006; Van Heuven, Dijkstra, & Grainger, 1998), as they are highly relevant to effects of inhibition and/or facilitation in bilingual language processing. Lastly, before moving on to the aims, predictions and methods of our study, I will explore the establishing of bilingual language profiles and second language (L2) proficiency through self-assessment and the degree of which they are reliable compared to objective standardized tests (Delgado, Guerrero, Goggin, & Ellis, 1999; Marian et al, 2007).

Effects of form similarity in first language word reading

Evidence shows that orthographic and/or phonological form similarity in single word reading affects the speed of word processing (e.g. Andrews & Lo, 2012; Davis & Lupker, 2006; Forster & Davis, 1984). One of the most commonly used techniques for investigating written word processing is masked priming, in which the subject is shown a prime word for a very short time – so short that the subject is unaware of its appearance, but long enough for the brain to process the information (usually 40-60 ms). Usually, the prime is shown in lower-case letters, followed by a symbol mask (e.g. "####"), while the target is shown in upper case (e.g. *tank-BANK*, *bank-BARK*, *kwqm-FISH*, etc). Masked priming tends to slow down processing of the target word when the prime is an orthographic neighbor (*corn-horn*), in which the prime and target differ by only a single letter in the same position in the word (Forster & Davis, 1984). This effect tends to increase when the prime word is of higher frequency than the target word (Davis & Lupker, 2006).

Words with large orthographic neighborhoods can be processed faster than words with small neighborhoods in lexical decision tasks, but a high-frequency orthographic neighbor or a high-frequency embedded word like "*car*" in "*scar*" can slow down responses lexical decision and naming tasks (Bowers, Davis, & Hanley, 2005). It is important to mention, however, that the very definition of an "orthographic neighborhood" contains variation, as it might be letter deletion (*last-blast*), transposed letters (*clam-calm*) and phonological overlap (*soup-hoop*) (Frisson et al, 2014).

Several models attempt to describe how words are recognized, such as the Multiple Read Out model (Grainger & Jacobs, Orthographic processing in visual word recognition: A multiple read-out model., 1996), the Dual Route Cascaded Model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), the Self Organizing Lexical Access and Recognition Model (Davis, 2001), and the Bi-modal Interactive Activation Model (Grainger & Holcomb, 2009). All of these models, are based on the Interactive Activation model (IAM) by McClelland and Rumelhart (1981), and agree that a letter string activates orthographically and/or phonologically similar words in the mental lexicon. If said string forms a word or a plausible non-word, activation increases over time, until it passes a threshold and a specific word is selected. As such, it is most relevant to our study to look at the core model: The Interactive Activation Model.

The IAM makes three core assumptions. First, it assumes that perceptual processing takes place within a system with processing at several levels, e.g. a letter level, a word level, a visual level, as well as "top-down" levels that provide input to the word level (McClelland & Rumelhart, 1981). Secondly, the IAM assumes that visual perception involves parallel processing: Both in that one is able to process at least four letters at the same time, and that processing at different levels can be simultaneous. Third, the IAM assumes that perception is an interactive process, where "top-down" and "bottom-up" processing works simultaneously and in conjunction in order to create multiple constraints that determine what we perceive. The Interactive Activation Model, therefore, suggests that the effects of lexicality and relative frequency are due to form priming involving both facilitation from sublexical overlap and inhibition from competition between lexical items. This invites a closer look at some of the experiments dealing with word processing. As the effects of proficiency is central to our study, we will begin by examining a study by Sally Andrews and Steson Lo (2012).

Andrews and Lo (2012) used the masked neighbor priming paradigm to examine the effects of proficiency on word recognition processes. They pointed out that most experimental psycholinguistic literature relies on average data for random samples of university students, which reflects a uniform impression of word processing among skilled readers. There is, however, individual variation among skilled reader samples on a variety of measures of written language proficiency. They tested university students assessed on reading, spelling and vocabulary skills. They investigated the effects of prime lexicality (where less priming is obtained because of form similarity when the prime is a word) for both transposed-letter (TL) primes (e.g. *sung SNUG; salb SLAB*) and neighbor primes (*snag SNUG; sleb SLAB*).

In previous experiments (e.g. Andrews & Hersch, 2010; Castles, Davis, Cavalot, & Forster, 2007), better spellers have shown inhibitory priming for higher frequency, orthographically similar words (e.g. *note NODE*), while poor spellers have shown facilitation for the same words. The inhibitory effect was restricted to words that are similar to, and thereby more easily confused with, many other words. Castles et al (2007) showed that better spellers showed significantly reduced facilitation from the ambiguous primes when unambiguous partial primes (e.g. *cr#wn CROWN* -as "*crown*" is the only possible word) were compared with ambiguous partial primes (e.g. *crow# CROWN* - where the word might be *crown*, *crowd*, *crows*). This suggests that they are sensitive to competition from alternative neighbors of the partial primes, which was not the case for poor spellers (Andrews & Lo, 2012). In other words, better spellers showed inhibition only in words where the neighbor primes had the potential to inhibit other lexical representations. Individual difference measures in the experiment by Andrews and Lo (2012) showed that shared variance in spelling, reading and vocabulary were associated with stronger inhibition from transposed-letter word primes and stronger facilitation from neighbor non-word primes. They explain this effect by pointing to the IAM, saying that competition between words plays an important role in lexical selection.

To summarize this section, evidence suggests that word form similarity affects word processing (e.g. Andrews & Lo, 2012; Davis & Lupker, 2006; Forster & Davis, 1984). The Interactive Activation Model (McClelland and Rumelhart, 1981) suggests that the structure of language processing and word form similarity might explain both facilitation from sublexical overlap and inhibition from competition between lexical items. These effects, however, are influenced by language proficiency and reading skill (Andrews & Lo, 2012; Bowers, Davis, & Hanley, 2005; Castles et al, 2007). As we have explored effects of proficiency and word form similarity in single word reading (Andrews & Lo, 2012), we now move on to exploring effects of proficiency in form similarity in sentence reading.

Form similarity in sentence reading: Effects of proficiency

In comparison to the literature on single word reading, very little research has investigated effect of form similarity in sentence reading. This is surprising, as very few of the words we read are encountered in isolation. Frisson et al. (2014) were interested in exploring how word reading processes work in normal sentence reading, as fluent reading requires constant activation and suppression of word candidates. Their research was inspired by Paterson, Liversedge and Davis (2009), who performed a silent sentence reading experiment where participants eye-movements were tracked while they read sentences like *"There was a blur as the blue lights of the police car whizzed down the street",* where "*blur*" is the prime and "*blue*" is the target. Results showed that the gaze duration on "*blue*" was significantly longer when preceded by an orthographic neighbor prime word (*blur*) compared to a control prime word

(*gasp*). In other words, high-frequency prime words (e.g. *wings*) slow responses to lower-frequency target words (e.g. *kings*), while neighboring non-words (e.g. *fings*) facilitate recognition of the same target word (Frisson et al, 2014).

Frisson et al (2014) argued that little attention has been given to the placement of the non-overlapping letter of two neighbors, which means that the mismatching letter can be in the word onset (*power-tower*), word offset (*turn-turf*) or mid-word (*axle-able*). This means that they can vary in the number of consonants, vowels or even syllables. If one assumes that phonology is inconsequential, this would not matter – but as research has shown, phonology does affect word reading (e.g. Frisson et al (2014).

Experiments using the phonological priming paradigm tend to show facilitation for end overlap items, while begin overlap items tend to show inhibition (e.g. Dufour, 2008; Titone & Connine, 1997). Experiments using the fast priming paradigm during reading, however, has found facilitation in begin overlap items. Because of this, Frisson et al (2014) chose to examine end-overlap and begin-overlap items separately. They performed two experiments, where they looked for evidence of form-based competition between words in sentences in English monolinguals. Experiment 1 examined whether different types of orthographic and/or phonological overlap resulted in different degrees of competition. Experiment 2 extended upon this, and was designed to examine whether the distance between the prime and target in the sentence structure affected the competition.

Experiment 1 showed that the inhibitory priming effect only applied to words that overlapped both at the orthographic and phonological level (*wings-kings, strain-strait*), while words that only overlapped at the orthographic (*bear-gear*) or phonological (*smile-aisle*) level had no inhibitory or facilitatory effect. They also discovered that inhibition only occurred in words that overlapped in the end (*pear-tear*), while words that overlapped in the beginning (*brat-brag*) caused facilitation. Frisson et al (2014) suggests that this may be because rhymes (end overlap) exerts a stronger phonological cue than begin overlap conditions do. However, they also pointed out that even though the findings of this experiment indicated that words must overlap in both orthography and phonology in sentence reading, this might not be the case in single-word studies. We will return to this issue in the discussion.

Experiment 2 showed that the distance between the prime/control and the target affects the inhibition effect, as it disappears when the distance is increased from 3 words to approximately 9 words. This agrees with the notion that activation levels of words decay quickly. This experiment also aimed to test priming in relation to episodic memory effects, which suggests that words are stored as episodic memory traces (including both orthographic and phonological features). When the same word was presented again, the memory trace would speed up processing. As such, when presented with a neighboring word, the memory trace might influence word identification and create an inhibition effect (Paterson, Liversedge, & Davis, 2009).

Motivated by earlier findings where reading skill affected the degree of inhibition (e.g. Andrews & Hersch, 2010; Andrews & Lo, 2012), Frisson et al (2014) also examined the difference in word processing between skilled and less skilled readers by dividing them into two groups: One containing skilled readers, while the other contained less skilled readers. Reading skill was measured using the Gray Silent Reading Test (Wiederholt & Blalock,

2000), which is a standardized test that measures reading comprehension. Participants read six short passages and answered short, multiple-choice questions related to each text.

In the short conditions of their study, the end overlap conditions showed *inhibition* for both skilled and less skilled readers, while the begin overlap conditions caused *facilitation*. Frisson et al (2014) theorized that this might be because rhymes (end overlap) exhibit a stronger phonological cue than begin overlap items. Furthermore, only skilled readers showed an inhibition effect for sentences in the long condition (where the prime and target were separated by 9 words), while less skilled monolingual readers were more likely to show facilitation. This indicates that the prime stayed active for longer in skilled readers. Significant effects were found only in the short lag conditions, which implies that there is a difference between skilled and less skilled readers in how they process the target when the distance between prime and target is longer. When the prime and target appeared in different sentences, however, there were no significant effects.

To summarize, reading skill affects word reading processes, where word form overlap priming causes more inhibition in skilled readers than in less skilled readers (Andrews & Hersch, 2010; Andrews & Lo, 2012). When word forms overlap in both orthography and phonology, only end overlap (rhymes) cause inhibition, while begin overlap items cause facilitation (Frisson et al, 2014). This raises the question of whether, and to what degree, these effects also apply to Norwegian-English bilinguals. In order to answer this question, however, we must examine some of the aspects of bilingual word processing and bilingual profile.

Reading processes in bilinguals

Models of non-selective language activation and suppression

As reviewed above, there is a great deal of evidence that reading processes in a first language involve a competition between words that overlap in orthographic and phonological structure. Interestingly, these competition processes relate to aspects of reading skill, which causes more skilled readers to slow down more than less skilled readers when reading neighboring words. To bilinguals, the processes might be even more complicated: The main hypothesis of bilingual language processing is that language activation is non-selective, which means that any input (heard or read) also activates the other languages that the subject knows. This means that it is impossible to suppress the inventory of any language completely, as the lexicon of all other languages known by the subject are activated automatically as soon as input occurs (e.g. Libben & Titone, 2009; Sunderman & Kroll, 2006). In other words, lexical competition in bilinguals can come from both (or all) languages they know.

When researching language non-selectivity, there are three main methods that tend to reoccur. All three mainly measure reaction time (RT) to determine the level of effect. The first of these is *word naming*, where the subject is shown an image or item and are asked to name it in the target language. Schwartz, Kroll and Diaz (2007) performed such an experiment with English-dominant English-Spanish bilinguals, where they examined cross-language interactions by asking whether the degree of phonological activation was affected by the degree of orthographic overlap. Their critical materials were cognates (words that are similar between the two languages in orthography, phonology and semantic meaning, e.g. *base* or *piano*). They aimed to explore this by presenting 18 English-Spanish bilinguals with a total of 240 Spanish words and their English translations, where half were cognates and half were

non-cognates. Participants named words in two blocks of trials, one English and one Spanish, and were told to name the words aloud, one at a time, and to do it quickly – if they were unsure about the pronunciation, they were told to guess. Each word appeared when the participant pressed a button on a keyboard, and reaction time was measured in milliseconds from the onset of stimulus presentation to the onset of articulation. When the naming tasks were done, the participants completed a language history questionnaire, where they rated their English- and Spanish proficiency in reading, writing, speaking and speech comprehension on a scale from 1 to 10. Results showed that when there was a mismatch between orthography, phonology and semantics, the mismatch caused delay in word naming (e.g. base: [beis] vs. ['ba.se]). As such, they suggested that the degree of activation might have been increased when phonological codes were highly similar, which would have produced activation of the orthographic representation from the other language, thereby inhibiting performance (Schwartz et al, 2007).

The second method in researching the selectivity of language is *lexical decision tasks*, in which a subject is shown a written word or pseudo-word (a word without meaning, but which obeys the orthographic rules of the language) and asked to identify whether it is a real word or not. Generalized lexical decision tasks ask if it is a word in any language, while language specific lexical decision tasks ask for word validity in the target language only. One such study was conducted by Lemhöfer and Dijkstra (2004) where they investigated how cross-linguistic overlap in semantics, orthography and phonology affected bilingual word recognition through four experiments. Their aim was to investigate the representation of cognates and interlingual homographs (words that look and/or sound similar, but mean different things) in the mental lexicon and to test the predictions of the BIA+ model (Dijkstra & Van Heuven, 2002) described below. They conducted both language-specific and general lexical decision tasks – some including interlingual homographs and/or interlingual homophones, while others included cognates. Experiment 1 was an English Lexical Decision experiment, where they tested 20 Dutch-English, native Dutch bilinguals in tasks involving words that overlapped in Orthography (O, e.g. *glad*), Phonology (P, e.g. English *cow vs* Dutch *kou,"*cold*"*) or both (OP, e.g. *spot*, English: /spɒt/ vs Dutch: /spɔt/, meaning *mockery* in Dutch). Participants read a written instruction in English before the experiment started, and did 30 practice trials before the test began, where they were presented with 180 trials, which in total included 15 items from each category (O, P, and OP). Experiment 2 was similar in participant group (20 native Dutch, Dutch-English bilinguals), tasks (English Lexical Decision) and procedure (30 practice trials, 180 trials, 15 overlapping items from each overlap category) with the central difference being that the words also overlapped in Semantics, meaning they were cognates instead of interlingual homographs/homophones. Experiment 3 and 4 had the same structure but had 34 participants in each and were focused on Generalized Lexical Decision. Experiment 3 included interlingual homographs, while Experiment 4 included cognates instead.

Lemhöfer and Dijkstra (2004) found that the time courses of word acceptance and non-word rejection differed between L1 and L2, and that when stimulus was presented, word candidates in both languages were activated, where L2 lexical codes were activated slower than L1 lexical codes. As such, cross-linguistic facilitation and inhibition effects were present either when L2 was the target language or when the item was a cognate. They also noted that interlingual homographs in generalized decision tasks did not produce cross-linguistic effects,

which indicated that response was based on the earliest available lexical code (which, with the participants in their experiment, was Dutch orthography).

The third of the main methods of language processing experiments is *word priming*, where subjects see a word (which is either "related" or "unrelated") before being presented with a target word, where the prime word is meant to produce inhibition (slowing down) or facilitation (speeding up) of the subject's reaction time. This method has already been explained in a previous chapter (e.g. Andrews & Lo, 2012; Davis & Lupker, 2006; Forster & Davis, 1984; Frisson et al, 2014), and as it is the primary method of our study, will receive further attention throughout this study.

The research reviewed above has resulted in several models of how language is processed in bilinguals, and that during word recognition, bilinguals activate word forms in both (or all) languages in word recognition. As such, even skilled bilinguals cannot deactivate either of the languages during processing, which applies in both directions – L1 affects L2 processing, and L2 affects L1 processing. To further examine bilingual word processing, we will examine three relevant models: The Bilingual Interactive Activation model (Van Heuven, Dijkstra, & Grainger, 1998), the Revised Hierarchical Model (Sunderman & Kroll, 2006) and the Inhibitory Control Model (Green, 1998).

Evidence for the BIA model

Van Heuven, Dijkstra and Grainger (1998) claimed that interlingual interference occurs in bilinguals at the phonological level (foreign accent), sentence level (borrowed syntax) and lexical intrusions from the other language (accidental lexical borrowing). They aimed to examine how word recognition was affected by existing word neighbors from other languages. To explore these processes, they performed a series of progressive demasking and lexical decision experiments.

Experiment 1 and 2 involved the word identification paradigm called Progressive Demasking, which alternates the presentation of a target word with that of a mask. Throughout the alternation, target presentation time is slowly increased, while that of the mask decreases.

In experiment 1, 42 Dutch students were divided into a High-Proficiency (HP) group and a Low-Proficiency (LP) group based on a questionnaire about language proficiency, completed after the trials of the experiment. In the trials, participants were presented with two blocks of items, one for each language, where each block consisted of 80 items (in addition to 25 practice trials for each block). They were placed in front of a computer screen and pressed a button to activate each trial. When they pressed the button, two small lines appeared above and below the center of the screen. After 1500ms, the screen cleared, and a checkerboard mask appeared at the center of the screen. After 300ms, the target word appeared for 15ms, followed by a checkerboard mask of the opposite pattern (e.g. black-white-black-white became white-black-white-black) which appeared for 285ms. The target word appeared for 30ms, and this process continued until mask presentation time was 0ms, until the participant pushed the response button, or until a time-out period of 6 seconds was reached. Van Heuven et al (1998) found that target word recognition was influenced by the number of orthographic neighbors in the non-target language, causing inhibition when the word had many neighbors (a significant effect in Dutch item- and participant analysis, while the effect of English neighbors only was significant in participant analysis) which cannot be explained without a

non-selective access model of language processing. They examined this effect further in their second experiment.

Experiment 2 used the same materials and task as Experiment 1 but tested 40 native Dutch Dutch-English bilinguals by presenting the stimuli in a single block, containing both English and Dutch words in random order. Results showed significant inhibition of increasing the number of neighbors for both English and Dutch items, where non-target language neighbors affected identification of targets in the progressive demasking task. For within-language neighborhoods, facilitation was observed in English target words, while Dutch words showed inhibition. These findings encouraged Van Heuven et al (1998) to examine whether the effects also extended to lexical decision tasks.

Experiment 3 was generalized lexical decision tasks, in which 48 participants native Dutch Dutch-English bilinguals determined whether a string of letters was an English or Dutch word (opposed to a non-word). The trials included the same stimuli as in Experiment 2, but with the addition of 4 non-word conditions, varying in the number of orthographic neighbors in Dutch and English. Non-word condition 1 had many neighbors in both languages (mean 3,5 for both). Condition 2 had many in English (mean 3,5) but few in Dutch (mean 1), condition 3 had few in English (mean 1) but many in Dutch (mean 3,5), and condition 4 had few (mean 1) in both languages. In the experiment, participants were told to decide as quickly as possible whether the presented item was a word (in English or Dutch) or not. Half of the participants were told to respond with their left finger for a Dutch or English word and with their right if it was a non-word, while the other half received opposite instructions. After 24 practice trials, the 320 experimental stimuli were presented in pseudo-random order, with no more than four words or non-words in sequence. Stimuli remained on the screen for 1500ms or until a response button was pressed (Van Heuven et al, 1998). Results showed that the effects from experiment 1 and 2 did indeed extend to lexical decision tasks, as Dutch neighbors caused significant inhibition on English target words, while within-language neighbors caused facilitation in English stimuli. L1 (Dutch) caused the most influence in the effects of rejection of non-words (both in English and Dutch). To further explore this, experiment 4 tested the same English words with a different group of Dutch-English bilinguals with a control group of English monolinguals.

Experiment 4 was language-specific lexical decision, where 20 English monolinguals and 21 native Dutch users were instructed to identify whether strings of letters were words in English or non-words as quickly as possible by pressing the right button if the string was an English word, or the left button if it was a non-word. They found that the responses of English monolinguals were not influenced by the number of Dutch neighbors, but it did influence Dutch-English bilinguals significantly, despite the fact that there was no Dutch stimuli. This strongly suggested that knowledge in the non-target language affects processing.

In order to explain their findings and the non-selectivity of language processing, Van Heuven et al (1998) proposed the Bilingual Interactive Activation (BIA) model (illustrated in Figure 2). This model claims that when a bilingual is presented with a string of letters, numerous lexical candidates in all languages compete for activation. Language nodes (criteria) will restrict the forms that are deemed incompatible with the target language, and competition and inhibition will gradually produce a "winner", while the remaining forms are suppressed. The hierarchy of processing in the BIA is shown in Figure 2.

The BIA model

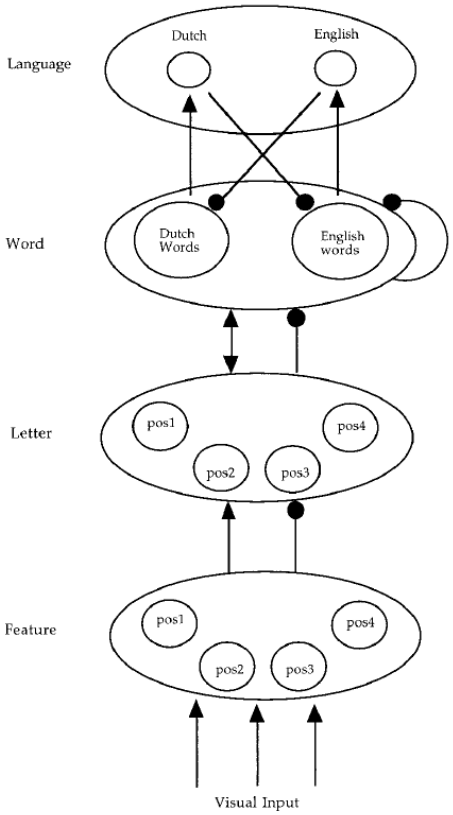


Figure 1: An illustration of the Bilingual Interactive Activation Model.

The process described in the BIA model (for visual input) places the feature level as the first. At this level, letters have not yet been identified, so the feature "-" might represent "A", "E", "H". When combined with the feature "I ", however, this causes less activation for "A", but more activation for the letters "E" and "H". When the feature " I" is added, "E" is excluded, which leaves "H" as the candidate with the highest level of activation. The same process is applied at the letter level, where positional information influences activation. If the letter "H" is the first letter, this causes more activation for every word the subject knows that begins with "H". When "H" is followed by "A", this causes activation of all words the subject knows that begins with "HA". Finally, with the addition of "T", the unit "HAT" can be formed. If this unit is a word the subject knows, it is activated as a word. If it is a word in both languages, however (Norwegian HAT= hate), the subject must decide which concept to link the word to. If the target language is English, the concept for "*headwear*" will be activated more than the Norwegian "*negative emotion*".

The BIA, then, implements both non-selective, bottom-up processing, where letters activate words from both languages in the mental lexicon, and language-specific, top-down processing, where language nodes selectively inhibit activity in the words of the other language. However, this model makes no claims about effects of language proficiency on these inhibitory processes. In contrast, the Revised Hierarchical Model (Sunderman & Kroll, 2006) does make claims about proficiency and bilingual language processing.

Evidence for the RHM model

The Revised Hierarchical Model (RHM) (Sunderman & Kroll, 2006) models the mapping of words-to-concepts. It claims that L1 words have a direct connection to conceptual meaning, while L2 (in the early stages of acquisition) depends on translation through L1 in order to access meaning. In other words, only advanced users of L2 are able to access conceptual meaning directly, without going through L1 first.

Sunderman and Kroll (2006) performed an experiment with bilinguals (L1 English, L2 Spanish) as an extension of a study by Talamas, Kroll and Dufour (1999) in order to test the predictions of the BIA model and the RHM. They aimed to examine (1) if lexical information in L1 was active in the processing of L2, (2) whether this activation was influenced by proficiency, (3) whether L2 lexical access is faster with increasing proficiency and (4) whether grammatical class functions as a cue to lexical access, and if so, if this is influenced by proficiency.

Sunderman and Kroll (2006) recruited 107 native English-speaking university students were recruited from Spanish language classes. Prior to the experiment, they completed a language history questionnaire where they rated their L1 and L2 reading, writing, speaking, and oral comprehension skills on a scale from 1 to 10. To assess cognitive abilities in the participants, they organized a Reading Span Task, in which participants decided whether 80 sentences (one at a time, divided into four sets of 2, 3, 4, 5 and 6 sentences) were semantically plausible. After a set, the participants were asked to recall the final word of each sentence in the set. Each trial presented a fixation point for 300ms, followed by a sentence that remained until the participant pressed a response button. When they had pressed the button, the participant wrote their response in a booklet, and pressed a key to continue to the next trial. The Reading Span Task was followed by a Picture-Naming Task, where participants were presented with 40 pictures of dictionary-like drawings and asked to name the picture aloud in Spanish as quickly as possible (after 10 practice trials).

For the actual experiment, Sunderman and Kroll (2006) generated 48 correct translation pairs (e.g. *cara-face*), each with 6 distracters – 2 for each of 3 conditions (same and different grammatical class). The first condition was an orthographic neighbor (orthographically similar) to the first item of the pair (e.g. *card-care*), the second condition was an orthographic neighbor to the second item of the pair (e.g. *fact-fast*), and the third condition was meaning related (e.g. *head-pretty*). Participants were tasked to decide as quickly as possible whether two words were translation equivalents and were presented with the L2 (Spanish) word followed by the L1 (English) word. Each trial was triggered by a button press, where the L2 word was presented for 400ms, followed by a 100ms blank screen before the second word appeared in the same position, and remained until the participant responded.

In their experiment, the BIA predicted that words such as *gato* and *gate* would produce competition in word recognition, thereby creating a longer response time. The RHM predicted that the nature of lexical activation was at the level of translation equivalents (*gato-cat*). The RHM also predicted that the response time would increase in tandem with L2 proficiency, where the skilled users had short response times, while the less skilled responded more slowly. Their findings showed that the predictions of both models were correct. Both proficient and less proficient L2 users showed activation of lexical form neighbors, but only the less skilled learners activated the L1 translation equivalent – which confirmed the predictions made by the RHM. They also found that interference was reduced or eliminated when the grammatical class of the two words differed, which neither the BIA nor the RHM took into account (however, this is included in the revised version of the BIA, the BIA+). These effects took place in both more and less proficient learners, which disagrees with the predictions of the RHM, as the less proficient group also proved to be sensitive to conceptual information in L2 processing (Sunderman & Kroll, 2006).

They concluded that both lexical form relatives and translation equivalents in the L1 influence performance, but that it was more prevalent in less proficient learners. In form-related lexical neighbors, inhibition occurred regardless of proficiency. Additionally, they concluded that access to conceptual meaning can be accessed through the L2 without always having to go "through" L1. This means that although some of the central predictions of both models were shown to be correct, neither model is able to fully describe the complexity of lexical competition.

The Inhibitory Control Model

Green (1998) took inspiration from the Revised Hierarchical Model to develop the Inhibitory Control Model (ICM or IC model), which had several levels of control in bilingual word processing. The IC model proposes that inhibition occurs at the lemma level (*run*, *runs*, *running* and *ran* are all lexemes, but *run* is the lemma). As lexical nodes carry language tags, the words that have tags corresponding to the non-target language are reactively inhibited, which eliminates them from selection (Green, 1998). As such, the conceptual system activates lexical nodes of both languages of the bilingual but suppresses the non-target language.

The ICM model has three important features: Firstly, inhibition is reactive. This means that the inhibition only functions *after* the lexical nodes have been activated. It also assumes that the level of activation correlates positively with inhibition, in the sense that lexical nodes with a higher degree of activation will cause more inhibition. Second, the lexical nodes of the non-target language will interfere during lexical selection in the target language. Third, the processing between lexical and sublexical levels is discrete, which means that only the selected lexical node has phonological activation.

Because the activation of lexical nodes correlates with the degree of suppression needed, the IC model predicts that if the bilingual is proficient in the non-target language, the level of suppression would cause more inhibition. In other words, bilinguals who are proficient in their L2 will have more inhibition than less proficient users of L2. As such, Green (1998) suggested that lexical concepts are language specific. This means that a Norwegian-English bilingual has two lexical concepts for any object. Lexicalization activates a lexical concept, which activates a lexical node. The selection includes a checking procedure, which checks if the lexical node corresponds with the intended lexical concept. Lexical inhibition is therefore unnecessary to select the correct lexical node, as the checking mechanism guarantees that the given lexical node corresponds to the intended language-specific lexical concept (Kroll & De Groot, 2005). It is worth mentioning, however, that although the Inhibitory Control Model assumes that language inhibition exists, the model can, because of the checking mechanism, account for bilingual language selection without needing language inhibition (Kroll & De Groot, 2005).

To summarize, we have reviewed some central models in bilingual word processing and suppression, language activation and language suppression. Evidence strongly suggests that language activation is non-selective, and that orthographic neighbors do affect word processing in single word reading tasks. Our study, however, is interested in how language processing and lexical neighborhoods are processed in sentence reading. As such, Libben and Titone (2009) offers some valuable insight into the subject.

L2 sentence reading and eye-tracking

Libben and Titone (2009) investigated language non-selectivity by recording eye movements of French-English bilinguals reading sentences which included cognates (e.g. *piano*), interlingual homographs (e.g. *coin* (which means "corner" in French) or matched control words. Each of the sentences provided either a low or a high semantic constraint for target-language meanings, exemplified in Table 1.

Table 1: Sample sentences across conditions for interlingual homographs and cognates in Libben and Titone (2009).

Sample sentences across conditions		
Word type	**Low-constraint sentence**	**High-constraint sentence**
Interlingual homographs		
Target word	Since they really liked each other, they had an extended *chat* that lasted all night.	Since they liked to gossip, they had an extended *chat* that lasted all night.
Matched control	Since he was kind of bored, he made an extended *tune* that was very catchy.	Since he liked to compose songs, he made an extended *tune* that was very catchy.
Cognates		
Target word	Because they owned a lot of property around the world, the expensive *divorce* was a disaster.	Because of the bitter custody battle over the kids, the expensive *divorce* was a disaster.
Matched control	Because her parenst strongly disapproved of her decision, the expensive *wedding* was a disaster.	Because the maid of honor and best man were late, the expensive *wedding* was a disaster.

They found that both early- (e.g. first fixation duration, gaze duration and skipping) and late-stage comprehension measures (e.g. go-past time and total reading time) showed significant facilitation for cognates and inhibition for interlingual homographs in low-constraint sentences, but that late-stage comprehension measures showed no evidence of either. As such, the results indicated that semantically biased contexts resolve lexical access rather quickly (Libben & Titone, 2009).

This provides further evidence of the non-selectivity of language activation. The finding that cognates cause facilitation suggests that lexical activation is reinforced when the word is the same (and has the same semantic meaning) in both languages, while the inhibition caused by interlingual homographs suggests that lexical competition occurs between L1 and L2 regardless of target language. To our study, this also informs us that stimuli sentences should be low-constraint, and that both cognates and interlingual homographs should be avoided in the target regions.

Bilingual profile and L2 proficiency

As mentioned before, this paper is interested in word neighbor priming effects in bilinguals, and how those priming effects are related to bilingual profile and L2 proficiency. As such, the bilingual profile of our participants was established using data attained through a revised version of the Language Experience and Proficiency Questionnaire (LEAP-Q). (Throughout this paper, the term "bilinguals" refers to speakers of more than one language. As such, it also encompasses the term "multilinguals".)

There is substantial evidence that self-assessment is a reliable way to acquire data about a subject's bilingual profile (e.g. Marian et al, 2007; Delgado, Guerrero, Goggin & Ellis, 1999). However, several aspects may influence performance in a situation of self-assessment in language proficiency.

Self-assessment of language proficiency is more accurate when the participant is asked to rate specific language skills, such as reading, writing, speaking, and listening (Delgado et al, 1999). In order to examine the accuracy of self-assessment tests, Delgado et al (1999) performed an experiment with 80 introductory psychology students, who were relatively balanced Spanish-English bilinguals. The participants performed a picture naming task, complete a biographical questionnaire (ethnicity, education and language background) and rate their language skills on 10 questions, where they rated their overall fluency, reading- , writing- , speaking-, and listening skills on a scale of 1 (very poor ability) to 5 (excellent ability) in everyday activities. For their objective tests of the participants' bilingual abilities, Delgado et al (1999) used the Spanish and English versions of the Woodcock-Muñoz Language Survey (Woodcock & Muñoz-Sandoval, 1993). The forms contained four subtests: Picture Vocabulary, Verbal Analogies, Word Identification and Dictation. The mean of the scores were taken as the general English or Spanish proficiency measure, oral language ability was assessed by the mean of the Picture Vocabulary and Verbal Analogies tests, whereas the mean of Word Identification and Dictation tests assessed reading-writing ability. Participants first performed a picture naming task for 20 pictures in both Spanish and English in order to determine the level of bilingualism and be qualified for participation. Then, they were assigned to a Hispanic examiner for individual testing, where all instructions and materials were in English. When they had completed the testing, they were given the Woodcock-Muñoz survey in English and Spanish (in random order). Finally, they received feedback on their performance and were asked to make a second assessment on all 10 self-rating scales (in English for half of the participants, in Spanish for the other half). Their results showed that Hispanic bilinguals who are competent in both languages are more accurate in assessing their overall fluency and specific language skills in L1 (Spanish) than in L2 (English). The use of feedback lowered the ratings for both L1 and L2, but had a larger effect on L1 ratings (Delgado et al, 1999).

Delgado et al (1999) attempted to explain this by saying that participants receive more practice and feedback on English reading and writing skills compared to their English oral skills. They also point out that the assessment questionnaire asked for an evaluation of skills in "everyday activity", and that there was no specific comparison group. As such, Delgado et al. (1999) reflect that they might have found a better match if they asked for a self-assessment of language use in academic settings.

Marian, Blumenfeld and Kaushanskaya (2007) pointed out that experiments with bilinguals often yielded inconsistent findings in lexical- and phonological/orthographic processing depending on the bilinguals' age of acquisition, history of use and degree of proficiency and dominance. To establish a bilingual language profile, proficiency ratings in themselves were insufficient, as language learning and language use experiences play a significant role in linguistic competence. Because of this, they wanted to develop a valid and reliable questionnaire for efficient assessment of the linguistic profiles of bilinguals, which resulted in the LEAP-Q (Marian et al, 2007). The LEAP-Q has several key points in assessing the language profile of a bilingual. First was *language competence*, which was divided into

language proficiency, language dominance and language preference. The language proficiency section included an assessment of literacy-oriented proficiency, grammatical proficiency, vocabulary knowledge and discourse abilities, as well as proficiency ratings in speaking, listening, reading and writing (Marian et al, 2007). The language dominance section measurement asked for an ordering of language dominance in both global (dominance) and specific (proficiency). The preference measurements comprised specific questions (e.g. preferred language when reading a text available in all languages) rather than general language preference.

Data was also collected on four aspects of language acquisition in relation to age: Age of initial language learning, age of attained fluency, age of initial reading and age of attained reading fluency. As the environment of learning also influences proficiency, the LEAP-Q also asked for years of formal education, years of residence in an L2 country, average self-estimated use of L1 and L2 and chronological age, as they all have been proven to influence age-of-acquisition effects on bilingual language dominance.

Finally, prior and current language exposure has been shown to influence sentence-level performance in bilinguals (Marian et al, 2007). Because of this, the LEAP-Q also asked for language exposure in a country, at school, at work and at home. However, current language use also influences proficiency, which made the LEAP-Q include questions about bilinguals' exposure to L1 and L2 in different situations, such as interaction with friends and family, exposure during reading, watching TV and listening to the radio, as well as self-instruction and language tapes.

Table 2: The thematic structure of the Language Experience And Assessment Questionnaire, listing the aspects of measurements taken in the questionnaire (adapted from Marian, Blumenfeld, & Kaushanskaya, 2007).

Language competence	Proficiency	Literacy-oriented proficiency Grammatical proficiency Vocabulary knowledge Discourse abilities Speaking proficiency Listening proficiency Reading proficiency Writing proficiency
	Dominance	Global (dominance) Specific (proficiency)
	Preference	Reading Writing Listening Speaking
Language acquisition	Age of acquisition	Age of initial language learning, Age of attained fluency Age of initial reading Age of attained reading fluency
	Learning environment	Years of formal education Years of residence in L2 country Average self-estimated use (L1 & L2) Chronological age
Language exposure	Location	In a country At school At work At home
	Situation	Interaction with friends/family Exposure through reading Exposure through watching TV Exposure through listening to radio Self-instruction (e.g. language tapes)

In order to assess the validity of the LEAP-Q, Marian et al (2007) performed two experiments. In the first, 52 participants representing 34 languages reported their own language history and proficiency using the LEAP-Questionnaire. Their answers were analyzed using both a factor analysis (8 factors, see table 2) and a multiple regression analysis, and became the base of a revised version of the questionnaire.

Table 3: Factors in the analysis of participant data of the first experiment (Marian, Blumenfeld and Kaushanskaya, 2007)

Factor 1	L1 Competence
Factor 2	Late L2 Learning
Factor 3	L2 Competence
Factor 4	L1 Maintenance
Factor 5	Late L2 Immersion
Factor 6	Media-Based Learning
Factor 7	Non-Native Status
Factor 8	Balanced Immersion

For the revised LEAP-Q, Marian et al (2007) had 50 Spanish-English bilinguals answer the questions, followed by a series of language tests to establish the degree of accuracy in the self-assessment of the questionnaire. Seven aspects of language proficiency were tested:

1. Reading fluency: Participants read as many sentences as possible within a 3-minute interval, and decided whether each sentence was true or false.
2. Passage comprehension: Participants read passages and supplied missing words.
3. Productive picture vocabulary: Participants name pictures.
4. Oral comprehension: Participants listen to passages and supply missing words.
5. Sound awareness: Participants complete a rhyming task, a sound deletion task, a sound substitution task and a sound reversal task.
6. Receptive vocabulary: Participants identify pictures in response to auditory instructions.
7. Grammaticality judgement: Participants read 50 sentences in English and 50 in Spanish, and judged whether they were grammatically correct.

Their correlation tests showed that there were strong positive correlations between the standardized behavioral measures (the seven tests mentioned above) and self-reported measures of understanding, speaking and reading L1 and L2. The performance of the participants on the sound awareness and receptive vocabulary did not relate to the self-reported L1 proficiency, but was significantly related to L2 proficiency.

The majority of the standardized measures correlated more strongly with self-reported measures of L2 proficiency than L1 (with the exception of the grammaticality judgement test). The highest correlation values in L2 self-reported proficiency and the standardized behavioral measures were passage comprehension and oral comprehension.

To summarize, Marian et al (2007) found a strong correlation between self-reported and objective standardized proficiency measures, suggesting that self-assessment is a reliable way of establishing a bilingual profile (Marian et al, 2007).

Language comparison

As we have discussed, the degree of activation and competition predicted by bilingual models of language processing is related to the degree of similarity between the words of the languages spoken (e.g. Green, 1998; Libben & Titone, 2009; Sunderman & Kroll, 2006; Van Heuven et al, 1998). Because we used native Norwegian Norwegian-English bilinguals in our experiment, the construction of our materials was informed by the relations between Norwegian and English. Both English and Norwegian can be traced back to their common "ancestor", Proto-Germanic (as illustrated below). As both languages stem from Germanic, the phonology, morphology, syntax and several words are the same for the two languages. They also tend to "borrow" words from each other: While English uses words such as "club" (klubba), "scathe" (skaða) and "maelstrom" (malstrøm), Norwegian has taken words like "date", "lunch", "bacon" and "caps" from English. To the issue at hand, however, it is more useful to look at cognates and bilingual homographs.

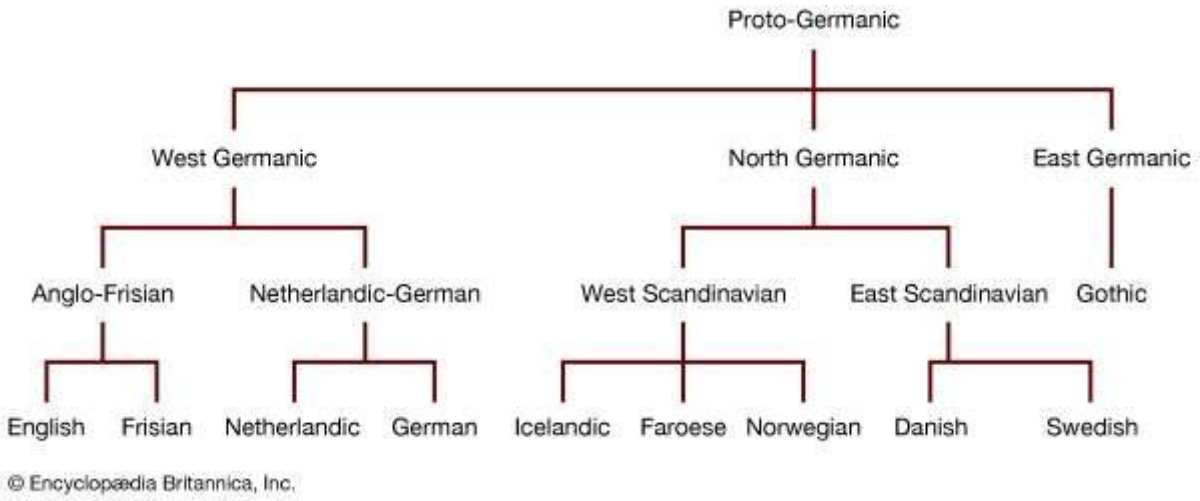
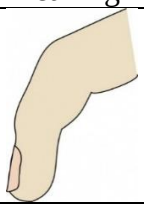
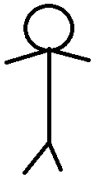
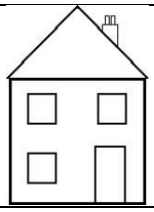


Figure 2: An overview of language family evolution/divergence from Proto-Germanic to the languages used today (Encyclopædia Britannica, 2017).

Cognates are words that are (close to) identical in orthography, phonology and semantics between two languages, while *bilingual homographs/homophones* are words where the orthographic/phonological form is the same, but the meaning is different. Examples of both categories and their relationship to semantic meaning is illustrated in Table 4. Cognates and bilingual homographs (also known as "false friends") affect reading speed and language processing, where cognates cause facilitation (speeding up) in reading, while interlingual homographs cause inhibition (slowing down) (e.g. Libben & Titone, 2009). Because of the common roots of the two languages, there are several examples of both cognates and interlingual homographs between Norwegian and English.

Table 4: An illustration of Cognates and Bilingual homographs/homophones in Norwegian and English, showing the relationship between orthography and semantic meaning. .

Cognates

Norwegian	English	Meaning
finger /ˈfɪŋər/	finger /ˈfɪŋgər/	
person /pæˈʂoːn/	person /ˈpɜrsən/	
hus /huːs/	house /haʊs/	

Bilingual homographs/homophones

Norwegian	Meaning	English	Meaning
full	drunk	full	no longer hungry
dress	suit	dress	clothing for women
bar	carried	bar	drinking place

Phonetic inventory

As mentioned earlier, Norwegian and English have common history, but have drifted apart. Most notably, this happened in two stages, which are commonly referred to as Grimm's Law (the first Germanic shift) and the second Germanic shift (Campbell, 2013). Grimm's Law described a sequential shift in phonetic inventory from Proto-Indo-European to Proto-Germanic, where voiceless stops became voiceless fricatives, voiced stops became voiceless stops, and voiced aspirated (murmured) stops became "plain" voiced stops. The Second Germanic Consonant Shift in West Germanic dialects between 300 and 900 A.D. The shift took place through three phases:

1. The Germanic voiceless plosives became fricatives in certain phonetic environments…
    a. English "ship" - /ʃɪp/
    b. Norwegian "skip" - /ʃiːp/
2. …while they became affricates in other environments:
    a. English "apple" - /ˈæpəl/
    b. Norwegian "eple" - /ɛplə/
3. The three voiced plosives became voiceless:
    a. English "door" - /dɔːr/
    b. Norwegian "dør" – /døːr/

(Campbell, 2013)

Old English was completely unaffected by this shift, while the other West Germanic languages (e.g. Norwegian, German) were affected (Hernández-Campjoy & Conde-Silvestre, 2012).

There is a difference in both the graphemes and the phonemes in English and Norwegian. While the English alphabet contains 6 vowels, they may represent 11 different phoneme monophthongs in British English. The same is true for Norwegian – the 9 vowels of the Norwegian alphabet (including *æ*, *ø* and *å*) can represent 18 different monophthongs in Eastern Norwegian. Figure 4 shows the 11 vowel monophthongs of British English next to the 18 vowel monophthongs of East Norwegian. As can be seen in the Figure 4, some sounds are exclusive to each language (e.g. /ə/ and /ʊ/ in English, /ʉ/ and /ø/ in Norwegian), and Norwegian has a larger number of front vowels.

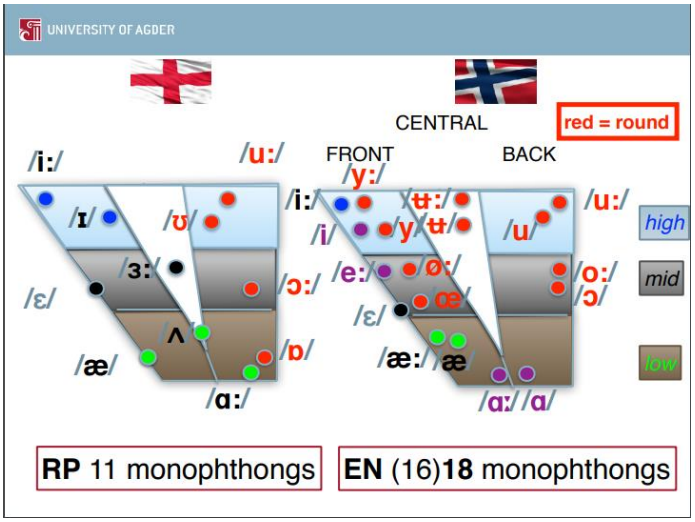


Figure 3: An illustration of the difference in the phonemic inventory of British English and East Norwegian.

(Wetterlin, 2017)

As with vowels, there is also a difference in the phonemic inventory of consonants between the two languages. As they each evolved into their respective contemporary forms, they both lost some sounds that used to belong to both languages, and gained some new variants of consonants.

*Table 5: A comparison of the consonant inventory of American English and East Norwegian. Common consonants are black. Exclusive consonants are marked in red for American English consonants, and blue for East Norwegian consonants. *The labio-velar /w/ is not included in the tableau, but is still considered a consonant in the American English phonemic inventory. (Kristoffersen, 2000)*

	Bi-labial	Labio-dental	Dental	Alveolar	Post-alveolar	Retro-flex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive/Stop	p b			t d				k g			ʔ
Nasal	m			n				ŋ			
Trill											
Tap/Flap				ɾ		ɽ					
Fricative		f v	θ ð	s z	ʃ ʒ	ʂ	ç		ʁ		h
Affricate					tʃ dʒ						
Lateral fricative											
Approximant		ʋ			r	ɹ	j				
Lateral approximant				l		ɭ					

Phonetic inventory comparison

In table 6, we show some situations in which Norwegian learners of English might have issues in pronunciation due to unfamiliar phonemes. As one of the aspects of our study is that the stimuli items (prime and target) need to have both orthographic and phonologic overlap, there is a chance that participants with low proficiency in pronunciation might show some signs of inhibition because the sounds are unfamiliar.

Table 6: An illustration of cases in which a Norwegian learner of English might encounter difficulties due to the differences in phonemic inventory between English and Norwegian (Wetterlin, 2017).

/θ//ð/	**th**ere is **th**e **th**ree **th**istle **th**eatre	/ðer ɪz ðə θriː ˈθɪs.əl ˈθiː.ə.ʈɚ/
/s//z/	**s**even **s**inging **z**ebras	/ˈsev.ən sɪŋɪŋ ˈziː.brəz/
/ʃ//ʒ/	mi**ss**ion & fi**ss**ion vs vi**s**ion & fu**s**ion	/ˈmɪʃ.ən & ˈfɪʃ.ən/ vs. /ˈvɪʒ.ən & ˈfjuː.ʒən/
/f//v//w/	**f**ine vs. **v**ine vs. **w**ine be **v**ery **w**eary of **V**ictor **W**alter **V**iceroy of **W**ales	/ faɪn/ vs /vaɪn/ vs /waɪn/ / bi ˈver.i ˈwɪr.i əv ˈvɪk.tɚ ˈwɔltɚ ˈvaɪs.rɔɪ əv weɪlz/
/ʧ//dʒ/	**j**u**dg**es go to **ch**ur**ch** to **j**ive	/dʒʌdʒɪz ɡoʊ ʈə ʧɝːʧ tə dʒaɪv/
/r/	**r**eally ve**r**y **r**osy	/ˈriː.ə.li ˈver.i ˈroʊ.zi/

Inflection

Word inflection, or word conjugation, refers to the different forms a word may take (or affixes/clitics may be attached) in order to follow the syntactic rules of a language, and may be one of the most challenging aspects to Norwegian speakers of English as a second language, as this in an area where the two languages differ in several aspects. Because the patterns of inflection in Norwegian and English is a rather large subject, we will restrict this discussion to the most common word classes in our stimuli set: Regular nouns and regular verbs. As can be seen in Table 7, Norwegian regular nouns have four different forms: indefinite singular, definite singular, indefinite plural and definite plural, each with their own inflectional morphemes. Each may belong to three separate genders, each with their own determiners in the forms of suffixes and enclitics in noun inflection. (For a more detailed discussion about the status of the definite article as enclitics and not as affixes, see (Lahiri, Wetterlin, & Jönsson-Steiner, 2005).

Table 7: Inflectional morphemes and enclitics in regular nouns in Norwegian.

	Singular		Plural	
	Indefinite	Definite	Indefinite	Definite
masculine	En gutt	gutt**en**	gutt**er**	gutt**ene**
feminine	Ei jente	jent**a**	jent**er**	jent**ene**
neuter	Et hus	hus**et**	hus	hus**ene**

The determiners in English, on the other hand, precedes the noun instead, as can be seen in Table 8. The determiner shows whether the noun is definite or indefinite, while plurality is denoted by the plural morpheme "-s" in regular nouns.

Table 8: Inflectional morphemes in regular nouns in English.

Singular		Plural	
Indefinite	Definite	Indefinite	Definite
a boy	**the** boy	boy**s**	**the** boy**s**
a car	**the** car	car**s**	**the** car**s**
an apple	**the** apple	apple**s**	**the** apple**s**

This difference in determinism and plurality causes some trouble for Norwegian L1 speakers as they are learning English as L2 – as less proficient readers might have difficulty identifying these aspects of a noun without stopping to think.

In contrast, verbs, pose a challenge to Norwegian learners of English because in the present tense, Norwegian only uses the ending "-er", while English uses a variety of forms depending on the subject of the sentence.

Table 9: A comparison of present tense verbs in Norwegian and English.

Norwegian	English	Norwegian	English
Jeg **er**	I **am**	Jeg hopp**er**	I jump
Du **er**	You **are**	Du hopp**er**	You jump
Han/hun/det **er**	He/she/it **is**	Han/hun/det hopp**er**	He/she/it jump**s**
Vi **er**	We **are**	Vi hopp**er**	We jump
Dere **er**	You **are**	Dere hopp**er**	You jump
De **er**	They **are**	De hopp**er**	They jump

This difference in verb conjugation might cause less skilled Norwegian readers of English trouble when trying to identify and assign meaning in verb phrases. (As such, for less proficient users of L2, this issue might have slowed down total reading time compared to what it might have been in L1.)

To summarize, the differences between Norwegian and English may cause some difficulty for native Norwegian users due to the change in phonemic inventory – especially in regard to the vowels and consonants that are exclusive or similar (but not identical) to either language. It is also possible that the conjugation of verbs and nouns may cause trouble for the less proficient participants of our study. Because cognates and interlingual homographs have a pronounced effect on word processing (e.g. Libben & Titone, 2009), we avoided both categories in our stimuli – both in the prime/control region and the target region.

This study – experiments, motivations, questions and predictions

As mentioned earlier, this study is a replication of Frisson et al (2014), who found that lexical neighbors and form similarity caused an effect in eye movements in monolinguals, where skilled readers showed inhibition effect from reading words with word-final orthography and phonology in sentences, while words with the same overlap in the word-initial position showed facilitation. Evidence suggests that this is due to form based lexical competition, as described earlier in the BIA model (Van Heuven et al, 1998), the RH Model (Sunderman & Kroll, 2006) and the IC model (Green, 1998). As more than half of the world's population is bilingual (Grosjean, 2010), this invites the question – do the same effects also occur in

Norwegian-English bilinguals? If so, are the conditions that produce these effects the same? Further, are these effects influenced by L2 proficiency?

We aim to test whether, and/or to what degree, words that overlap in both the orthographic and phonological level show the same effect in bilingual subjects. Furthermore, we also examine the effects of the distance between the prime/control and target words in the same group, and relate this to the subjects' bilingual profile established in the revised LEAP-Q.

Based on the research we have looked at so far in this paper, we predict that the results will be similar to those of Frisson et al (2014). There should be an inhibition effect for end overlap items and a facilitation effect for the begin overlap items in the short conditions, while the long conditions should show inhibition for skilled readers only, while less skilled readers should show facilitation.

The BIA model, however, would predict that the begin overlap items might cause inhibition due to lexical competition, while the RHM predicts that inhibition would be larger for less skilled readers, as they need to use translation in order to access conceptual meaning. The IC model predicts that we will see more inhibition in participants that are proficient in their L2 (English), and less inhibition in less proficient users of L2. As such, when we correlate the results with the bilingual profile of the subject, we expect that there are significant correlations between self-assessed reading skill and performance – especially that proficient L2 readers will show more inhibition in primed conditions than less proficient L2 readers.

Key manipulations

The stimuli consisted of 128 English sentences, which occurred in one of 4 different conditions (32 in each condition). In addition, each sentence could occur with target word preceded by a form-related prime or a control prime unrelated in form. This means that each participant encountered eight different forms of stimuli in total (see Table 10 for examples). Our study contains three main manipulations: The first is Lag Distance, which is divided into Short and Long conditions. In the Short conditions, the prime/matched control word have an average distance of 3 words from the target. The second manipulation is Relatedness, which is Related (prime word with overlap) or Unrelated (matched control word, no overlap). The final manipulation is Overlap Type, which consists of Begin Overlap (only the last grapheme/phoneme is different, e.g. *rail-raid*) or End Overlap (rhymes – only the first grapheme/phoneme is different, e.g. *jump-pump*). To exemplify, *Short End Related* refers to a stimuli item that has a Lag of 3 words (between the prime/control word) and has Overlap in the non-initial part of the word. A *Long Begin Unrelated,* however, has a Lag of 9 words, but does not have Overlap – instead of the prime, it contains the control word for a sentence where the prime would overlap with the non-final part of the target.

Method

Materials

Table 10: Examples of the different conditions and items in our experiment, showing both Lag length (short/long), Relatedness, Overlap type, Prime/Control words, target words and examples of sentences.

Condition		Prime/ Control	Target	Sentence
SHORT END	related	fork	pork	William used a big ***fork*** to eat the ***pork*** they had bought.
SHORT END	unrelated	knife	pork	William used a big ***knife*** to eat the ***pork*** they had bought.
SHORT BEGIN	related	seven	sever	He counted up to ***seven*** knights that could ***sever*** the dragon´s head.
SHORT BEGIN	unrelated	eight	sever	He counted up to ***eight*** knights that could ***sever*** the dragon´s head.
LONG END	related	wave	cave	Unfortunately a big ***wave*** forced them back out from the secret and unexplored ***cave*** on the beach.
LONG END	unrelated	tide	cave	Unfortunately a big ***tide*** forced them back out from the secret and unexplored ***cave*** on the beach.
LONG BEGIN	related	paint	pains	Drinking the old ***paint*** caused Leif to experience lasting and severe stomach ***pains*** and he was rushed to the hospital.
LONG BEGIN	unrelated	juice	pains	Drinking the old ***juice*** caused Leif to experience lasting and severe stomach ***pains*** and he was rushed to the hospital.

The related primes and targets were orthographic and phonological neighbors (both orthographically and phonetically similar, differing in a single phoneme/grapheme per item in the same position in the word). The differing grapheme/phoneme is the first unit in the "end" overlap condition (*fork/pork*), while it occurs at the very end of the word in «begin» condition words (*rail/raid*) (see table 9 for examples). Each word had a maximum of two syllables, and none of the prime/target/control words were cognates or bilingual homographs/homophones in English/Norwegian. The short begin and end groups had between 3 and 4 words separating the prime/control from the target, while the long begin and end groups contained 8 to 9 words between them.

The primes, targets and control words were matched for mean frequency in the sets, and the number of letters and phonemes are matched for each set. We allowed for a difference of up to 10 occurrences per million words in English in primes and control words, while the

maximum allowed difference in targets was set to 1. The target was always lower in frequency than the prime/control word.

No targets, primes or control words ever occurred more than once in any category. None of them were cognates, homographs or homophones. Every control and prime word was also as close to the same semantic meaning as possible, although frequency took precedent over semantics. They were, however, all grammatically correct and semantically plausible. Care was also taken to make sure that none of the control words had the same grapheme and/or phoneme in the same position as in its target (which means that *prime-coins* would not work, as "*i*" is the third letter in both words).

The sentences was constructed to be felicitous, non-ambiguous, grammatically correct, plausible and to have a low degree of constraint – that is, the target word should not be easily predictable. The full set of stimuli is shown in Appendix I.

Table 11: Mean frequencies and number of graphemes in control words, prime words and target words, showing the relationship between frequency and number of graphemes between short/long sentences and Overlap type.

	Control		Prime		Target	
Short end	characters	frequency	characters	frequency	characters	frequency
	4,781	69,643	4,781	62,053	4,781	12,363
Short begin	characters	frequency	characters	frequency	characters	frequency
	4,344	62,728	4,344	69,868	4,344	11,910
Long end	characters	frequency	characters	frequency	characters	frequency
	4,781	64,784	4,781	64,471	4,781	12,888
Long begin	characters	frequency	characters	frequency	characters	frequency
	4,656	64,601	4,656	65,127	4,656	12,556

Design

The main manipulations of the experiment was Overlap Type (Begin/End), Priming (Related/Unrelated) and Lag (Short/Long). All conditions occurred within each subject. For items, only Overlap Type occurred between items. Each subject saw half of their sentences in Related form and the other half in Unrelated condition, which alternated between participants.

The stimuli were distributed into eight pseudo-random blocks, where each subject was exposed to four (with a small break in between each). Every subject saw the same number of Related/Unrelated sentences, and never more than two Related sentences in a row. Each block contained an equal number of sentences from each condition. In order to ensure an evenly distributed set of data, the blocks rotated to make sure that any question would not appear in the same set for each participant.

Apparatus

Eye movement data were acquired using an SR Research Eyelink 1000 Plus Camera. Sample settings were set at 500hz. Viewing was binocular, but only data from the right eye was recorded. A target sticker was used in addition to a head rest for the subject to improve stability of the recording.

Procedure

All communication prior to the completion of the task was conducted in English, in order to prime the subjects for use of the language. (Research on Chinese-Taiwanese bilinguals has shown that language performance is affected by language switching in certain situations – see Wu & Thierry, 2010). Each subject was guided to the location, where they were given a set of written instructions on how to perform the experiment. The subject was then placed in front of the display screen and eye tracker, adjusted the seat to a comfortable position, and the experimenter repeated the most imperative practical instructions again (such as where the sentence appears, which buttons to press, etc.).

Each of the blocks began with five practice sentences in order to familiarize the subject with the controls for the experiment. The blocks consisted of 32 sentences, and each block was separated by a screen saying "PAUSE", in which the subject was encouraged to take a small break. When displayed on the screen, the control and target would always appear on the same line, as line skipping might affect gaze duration. This was accomplished by using the font COURIER NEW, where every character is the same height and width.

The participants were placed in front of a screen, with a Norwegian QWERTY-keyboard in front of them. The eye tracking camera was placed beneath the computer screen. Each participant was asked to adjust the chair to make sure they were sitting comfortably, after which the experimenter instructed them on how to control the computer during the experiment (where the sentence would appear, which keys to use, etc.) Complete instructions are included in Appendix II, while instructions for the experimenter are in Appendix III. The sentence appeared in the middle of the screen. When the subject had read the sentence and was confident they had understood its content, they were instructed to look at a grey circle in the lower right corner of the screen, and while looking at it, press the space bar. This triggered either a question (25% of sentences) or a statement saying "NO QUESTION", where they were instructed to answer "YES" to familiarize them with the location of the answer keys. When the question was answered, a calibration point was displayed on the screen, to help the experimenter determine whether there was any need for recalibration. If the difference was more than .5, a recalibration was performed until the difference was less than .5. The entire experiment (including welcoming the participant and asking them to read the instructions) took between 25 and 45 minutes. Finally, we only spoke Norwegian to the subjects when the experiment was done. This contributed to make sure that the subjects were indeed Norwegian-English bilinguals.

Participants

The participants were between the age of 18 and 35. They were all Norwegian native speakers, had English as a second language, and had no other home languages than Norwegian or English. They had no impairments in language (stuttering, dyslexia, etc.) or listening, and all participants had normal or corrected-to-normal vision (with glasses or

contact lenses). Most of the participants were undergraduate students at the University of Agder.

The participants completed a survey, which first determined eligibility (according to the requirements above), followed by the revised LEAP-Questionnaire. Eligible participants were then contacted for further participation. Every eligible person was offered a 200 NOK gift card when they had completed the survey and two experiments, including a language proficiency test and the eye tracking experiment. 31 participants were tested.

I will first report the results of the eye-tracking experiment and discuss them, followed by the results of the correlations with individual differences and a discussion of these findings.

Results: Eye-tracking experiment

In the eye-tracking experiment, three measures were analyzed: First fixation (early measure; the duration of the first fixation in the region of interest before leaving it), regression path (intermediate measure; the sum of all fixations from the time of entering the region of interest until going past it) and total time (late measure; sum of all fixations in the region of interest). There were two regions of interest: Target (the word where the results would manifest, including the preceding article belonging to the word) and spill-over (which was the next word if it was 4 characters or longer, in which case it included an additional word). The reading data were analyzed using a 2 (Overlap Type: end vs. begin overlap) x 2 (Relatedness: orthographically and phonologically related to the prime vs. unrelated) Analysis of Variance (ANOVA). Interactions were examined using means comparisons (t-tests). In addition, the eye movement data were compared to the data from the revised LEAP-Q using correlations.

Example:

	prime	*[control]*		*target*	*spillover*	
I was totally about to	*faint*	*[crash]*	when I saw	*a saint*	*walking*	on the water.

Short conditions

Table 12: Mean values of measurements from Short Lag conditions, including First Fixation, Regression-path, and Total Time measures, from Target and Spill-over regions, including both Overlap Type and Relatedness. All reading times are in milliseconds.

	Target				Spill-over			
	End overlap		Begin overlap		End overlap		Begin overlap	
	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated	Related
First fixation	275	276	269	286	262	274	275	262
Regression-path	408	435	469	501	428	484	487	485
Total Time	509	503	592	607	464	487	522	498

First fixation

Target region: In the target region, we observed an interaction (borderline in the items' analysis) between Overlap Type and Relatedness ($F_s(1, 30) = 4.7$, $p < .05$; $F_i(1, 62) = 3.0$, $p < .09$). Means comparisons showed no difference in first fixation duration for end-overlap items ($t_{s,i} < 1$), but a significant difference for begin-overlap items ($t_s(30) = 2.4$, $p < .05$, $t_i(31) = 2.1$,

$p < .05$), with longer fixation times when the target was related to the prime word. This means that **inhibition** only occurred for begin-overlap items in the earliest measure.

Spill-over region: The interaction between Overlap Type and Relatedness was also observed in the spill-over region ($F_s(1, 30) = 8.0$, $p < .01$; $F_i(1, 62) = 6.4$, $p < .05$). Means comparisons showed a borderline difference (**inhibition**) for end-overlap items ($t_s(30) = 2.0$, $p < .06$, $t_i(31) = 1.7$, $p = .10$) and a borderline **facilitation** for begin-overlap items ($t_s(30) = 1.9$, $p < .07$, $t_i(31) = 1.9$, $p < .07$). This suggests that there was a slight delay in the inhibition effect for the end-overlap items, although one should be careful interpreting this effect, as it did not reoccur in the other measures.

Regression-path

Target region: The regression-path analysis for the target region showed a main effect of Overlap Type ($F_s(1, 30) = 15.8$, $p < .001$; $F_i(1, 62) = 4.1$, $p < .05$).

The main effect of Relatedness ($F_s(1, 30) = 7.0$, $p < .05$; $F_i(1, 62) = 3.6$, $p < .07$), however, showed **inhibition** (longer reading times) in end-overlap items, which is in agreement with the findings of the original paper (Frisson et al, 2014).

Spill-over region: The spill-over region contained a borderline effect of Relatedness ($F_s(1, 30) = 3.3$, $p < .08$; $F_i(1, 62) = 1.9$, $p < .18$). The interaction approached significance in the subjects' analysis ($F_s(1, 30) = 3.4$, $p < .08$; $F_i(1, 62) = 1.8$, $p < .18$). When analyzing the the overlap conditions separately, we found inhibition for end overlap items ($t_s(30) = 3.1$, $p < .01$, $t_i(31) = 1.9$, $p < .07$), which was absent for begin overlap items ($t_{s,i} < 1$).

Total time

Target region: For the target region, Total Time analysis showed a main effect of Overlap Type ($F_s(1, 30) = 44.4$, $p < .001$; $F_i(1, 62) = 4.1$, $p < .05$), with end overlap items being read significantly faster than begin overlap items (506ms vs. 600ms).

Spill-over region: A similar effect of Overlap Type was observed for the spill-over region, though it did not reach significance in the items' analysis ($F_s(1, 30) = 7.2$, $p < .05$; $F_i(1, 62) < 1$). Again, the end overlap conditions were read faster than the begin overlap items (476ms vs. 510ms).

Long conditions

Table 13: Table 11: Mean values of measurements from Long Lag conditions, including First Fixation, Regression-path, and Total Time measures, from Target and Spill-over regions, including both Overlap Type and Relatedness. All reading times are in milliseconds.

	Target				Spill-over			
	End overlap		Begin overlap		End overlap		Begin overlap	
	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated	Related
First fixation	273	267	266	262	272	273	263	266
Regression-path	383	409	394	383	440	453	402	434
Total Time	490	497	466	437	494	468	458	467

No significant or near-significant effects were observed for the first fixation and regression-path measures.

Total time

Target region: Total time reading of the target region in the long conditions showed a main effect of Overlap Type in the subjects' analysis ($F_s(1, 30) = 7.2$, $p < .05$; $F_i(1, 62) = 1.4$, $p < .24$), with end overlap items taking longer to read than begin overlap items (494ms vs. 452ms).

Spill-over region: For the spill-over region, there was a borderline effect of Overlap Type in the subjects' analysis: ($F_s(1, 30) = 3.2$, $p < .09$; $F_i(1, 62) < 1$), again with end overlap items taking longer to read than begin overlap items (481ms vs. 463ms).

Discussion: Eye-tracking experiment

The results show that Norwegian-English bilinguals do show inhibitory effects for words following primes with orthographic and phonological overlap, and that there are some interesting correlations between the inhibition and L2 proficiency.

In the first fixation measure of the short conditions, the begin-overlap items caused inhibition, while end-overlap items showed no significant effect. In the spillover region, however, there was borderline effects of inhibition for end-overlap items, while begin-overlap showed borderline effects of facilitation. The short condition regression-path analysis for the target region showed a main effect of Overlap Type, but because the region was not controlled for length or frequency between the two overlap conditions, it is difficult to draw conclusions from this effect. However, the main effect of relatedness showed inhibition for end-overlap items in the target region, which was also found in Frisson et al (2014). Total time analysis in the short condition for the target region showed that end-overlap items were read significantly faster than begin overlap items. For the spill-over region, the effect was similar, but did not reach significance (although it might reach significance in further testing).

For the long condition, only the total time measures yielded significant or near-significant results. For the target region, end-overlap items took significantly longer to read than begin-overlap items. In the spill-over region, there was a near-significant difference.

In the principal components analysis, there were clear indications that L2 proficiency affects reading processes in bilinguals in different ways, and that the degree of proficiency can indicate how inhibition and/or facilitation effects manifest compared to "native-like" performance.

Our experiments aimed to examine whether, and to what degree, words with phonological and orthographic overlap showed the same effects in bilinguals as in monolinguals. We also wanted to explore the effect of distance on inhibition/facilitation effects, and to relate the results to a bilingual profile established through a revised version of the Language Experience And Proficiency Questionnaire (LEAP-Q).

We predicted that the results would be similar to that of Frisson et al (2014): There would be an inhibition effect for end-overlap items and a facilitation effect for begin-overlap items in the short conditions. The long conditions only would show inhibition for skilled readers, while less skilled readers were predicted to show facilitation for the long conditions. We did, however, prepare for the possibility that the begin-overlap items might cause inhibition due to lexical competition (the BIA model, see Van Heuven et al, 1998). On the other hand, the RHM indicated that inhibition would be larger for less skilled readers, as L2 access needs to go through L1 in order to retrieve the concepts behind lexical items (Sunderman & Kroll,

2006). The IC model predicted that L2 proficiency would correlate positively with inhibition, in the sense that we would find more inhibition in highly proficient L2 users than in less skilled L2 users (Green, 1998). From the results, it is clear that word neighbors (words which only differ by a single phoneme/grapheme) affect reading processes in Norwegian-English bilinguals in a similar manner to that of English monolinguals. It is also clear that the distance between the prime and target affect the size of the inhibition/facilitation effect – there were significant effects in several measures in the short conditions, but only the total time measure yielded significant inhibition for the long conditions.

The inhibition observed in the first fixation measures of the begin-overlap items in the short conditions agreed with the findings of Frisson et al (2014) with monolinguals, and may be explained by the IAM model. The model says that the letters in a word allow several "word candidates" (potential words) to remain active until the last letter. For example, the five-letter word "stee#" could be "steel", "steep", "steer" or "steek" (as long as the sentence is ambiguous), and the "winner" is only selected when the last letter is selected.

Although the effects did not reach the significance threshold of .05 in the t-tests, the fact that the first fixation measures of the spillover region showed a slight delay in the inhibition effect for the end-overlap items (compared to the facilitation of begin-overlap) is very interesting, as it was the opposite of what was found in monolinguals. Instead, Frisson et al (2014) found a slightly earlier effect of end-overlap (rhymes) compared to begin-overlap. This may be related to phonology, as rhymes exert a stronger phonological cue than begin-overlap – although further analyses are required in order to say anything definite about this difference.

As mentioned before, the main effect of Overlap Type shown in the regression-path analysis for the short condition items cannot provide much in regards to certainty, as the region was not controlled for length of frequency between begin- and end-overlap. The main effect of relatedness for the end-overlap items in the target region, however, agrees with the findings of Frisson et al (2014) in monolinguals. This effect may be explained through the BIA, in that several features comply with what the brain is "familiarized with" (primed for), but some features make that interpretation impossible, forcing the brain to process the word again. This also explains the total time analysis for the short conditions, in which end-overlap items were read faster than begin overlap items (significantly for the target region, borderline-significant in the spill-over region).

Interestingly, only the total time measure gave any significant (target region) or near-significant (spill-over region) results for the long conditions, where total reading time was longer for end-overlap conditions than for begin-overlap conditions. There were no significant or near-significant inhibition or facilitation effects in the first fixation or regression-path measures. The fact that the priming effect is present when prime and target is separated by 3 words, but disappears when separated by 9, may be explained by the general assumption that word level activation decays quickly (Paterson et al, 2009).

Frisson et al (2014) found that as long as the prime and target were in the same sentence, better comprehenders still showed an inhibition effect in the long condition items, while less skilled comprehenders did not. There are several possible explanations for this phenomenon. First, skilled comprehenders might retain superficial information for longer than less skilled comprehenders do. Second, it is also possible that the prime stays active for longer, as some evidence suggests that skilled readers rely more on phonological cues than less skilled

readers. Third, better comprehenders might have better memory than less skilled readers (Frisson et al, 2014). We will discuss this further below.

Results: Correlations with Principal Components of bilingual profile

Thirty subjects were included in the PCA correlation (with the results from the revised LEAP-Questionnaire). The data from one participant was deleted because the data set was incomplete. In the correlations, the difference for the unrelated and related conditions for each participant was calculated. Unrelated was deducted from the related, so that where a correlation is positive, the difference would indicate the level of inhibition. For example: if related = 600ms and unrelated = 500ms, the inhibition effect is (600-500=) 100ms.

The PCA component factors were based on the revised LEAP-Q from the related sister project (Handeland & Zitong, 2018), and were divided into 8 factors.

Factor 1 was associated with *L2 text proficiency*. The variables included chronological age, L2 exposure through reading and through interaction with friends and family, as well as proficiency in spelling, grammar, speaking and understanding.

Factor 2 was associated with *L1 proficiency*, and included variables like L1 exposure with friends, L1 regional accent, L1 learning through reading and TV, L1 proficiency in grammar, reading, spelling and writing, and years of formal education.

Factor 3 was related to *L2 speech proficiency*. It included self-assessed measures of L2 proficiency in pronunciation, speaking, vocabulary and grammar, as well as time spent in an L2 country. In addition, other variables included self-perceived accent in the participants' L1 and the importance of their L1 dialect.

Factor 4 was related to *L1 text proficiency*, where variables included L1 reading, spelling, speaking, writing, understanding and grammar, as well as years of education. This factor also included the age at which the participant started reading L2, the percentage of time they use L2 and how often others can identify their region of origin based on their L1 dialect.

Factor 5 was related to several aspects of both L1 and L2 proficiency, and included grammar- and writing proficiency, language learning and language exposure of both languages. The factors also included the age where the subject began reading L2. For simplicity's sake, this factor may be described as *general language proficiency*.

Factor 6 was *L1 learning*, and mostly concerned itself with the mediums through which learning took place, e.g. music, TV, reading and school.

Factor 7 was *L2 speech learning*, which included exposure and learning mediums. It also contained variables of language intrusion (both L1 into L2 and L2 into L1, the age when the participant began reading L1 and how important the subject thought it was to have a native-like L2 accent.

Factor 8, the last factor, was the *L2 age of acquisition*. It included self-assessed measures of ages where the participant began speaking and reading L2 fluently, age where they began acquiring and reading L2, grammar proficiency, learning in school (both for L1 and L2) and the importance of dialect in L1 and accent in L2. (The complete table of factors and variables can be found in Appendix IV.)

Correlations for the short conditions

For the short conditions, there were no significant correlations with first fixation target- or spillover region, nor for the regression-path target- or spillover region. Total time measurements of the target region, however, showed a negative correlation for the end-overlap items and factor 5 of the LEAP-Q PCA analysis (L1 & L2 language proficiency, reading and writing skills). There was also a strong correlation for begin-overlap items and factor 1 (L2 text proficiency), which means that the more proficient the subject is in L2, the larger the inhibition effect. The total time measure for the spill-over region had a negative correlation for end-overlap items and factor 2 (L1 proficiency). This means that the better the subject is in their native language, the smaller the inhibition effect.

There were also 3 positive correlations for begin-overlap items and factors 1 (L2 text proficiency), 2 (L1 proficiency) and 5 (L1 & L2 language proficiency, reading and writing skills). In other words, the more proficient the subject was in each of the factors, the larger the inhibition effect.

Table 14: Correlations and significance measurements for Principal Components Analysis of the Short Lag condition. All significant or near-significant effects are marked in yellow.
***. Correlation is significant at the 0.01 level (2-tailed).*
**. Correlation is significant at the 0.05 level (2-tailed).*

		L2 text proficiency	L1 proficiency	L2 speech proficiency	L1 text proficiency	General language proficiency	L1 learning	L2 speech learning	L2 age of acquisition
SHORT_TT_endoverlap_target	Pearson Correlation	-,168	-,195	-,052	-,102	-.394*	-,169	-,147	-,278
	Sig. (2-tailed)	,375	,301	,786	,592	,031	,372	,439	,137
SHORT_TT_beginoverlap_target	Pearson Correlation	.548**	,106	,238	-,122	,239	,229	.368*	-,076
	Sig. (2-tailed)	,002	,578	,205	,522	,204	,223	,045	,691
SHORT_TT_endoverlap_spill	Pearson Correlation	-,137	-.384*	-,017	-,226	-,192	-,326	-,242	-,120
	Sig. (2-tailed)	,469	,036	,931	,229	,311	,079	,197	,528
SHORT_TT_beginoverlap_spill	Pearson Correlation	.584**	.492**	,215	,236	.539**	,273	,226	-,028
	Sig. (2-tailed)	,001	,006	,255	,209	,002	,145	,231	,882

Correlations for the long conditions

For the long conditions, there were some correlations that were difficult to interpret. Sometimes, subjects with a higher degree of L2 proficiency showed smaller inhibition effects. However, this might be an effect of working memory due to the longer distance between prime/control and target.

Table 15: Correlations and significance measurements for Principal Components Analysis of the Long Lag condition. All significant or near-significant effects are marked in yellow.
***. Correlation is significant at the 0.01 level (2-tailed).*
**. Correlation is significant at the 0.05 level (2-tailed).*

		L2 text proficiency	L1 proficiency	L2 speech proficiency	L1 text proficiency	General language proficiency	L1 learning	L2 speech learning	L2 age of acquisition
LONG_FF_endoverlap_target	Pearson Correlation	.425*	.491**	,064	,102	,283	,309	,123	-,134
	Sig. (2-tailed)	,019	,006	,738	,592	,129	,096	,517	,479
LONG_RP_endoverlap_target	Pearson Correlation	-,036	-,254	,085	,072	-,236	-.456*	,061	-,083
	Sig. (2-tailed)	,851	,175	,654	,705	,210	,011	,750	,663
LONG_RP_beginoverlap_target	Pearson Correlation	-.677**	-,199	-.426*	-,080	-.364*	-,204	-,176	.445*
	Sig. (2-tailed)	,000	,291	,019	,673	,048	,279	,351	,014
LONG_RP_endoverlap_spill	Pearson Correlation	-,221	-,034	-,057	-,066	-,269	,123	,106	,121
	Sig. (2-tailed)	,240	,858	,765	,730	,150	,516	,577	,523
LONG_RP_beginoverlap_spill	Pearson Correlation	,197	.381*	-,220	,012	.388*	,101	,259	-,001
	Sig. (2-tailed)	,296	,038	,242	,950	,034	,596	,166	,997

Discussion: Correlations with PCA factors

The second aim of our study was to relate priming effects to individual differences in language profile and proficiency. There were some significant correlations between bilingual profile and the inhibition effects. The negative correlation of end-overlap and the factor of *General language proficiency* and the strong correlation between begin-overlap and the factor of *Text proficiency* was expected. The more proficient the bilingual is in their L2, the more native-like the inhibition effect is. This agrees with the RHM model, which says that access to L2 must go through L1 until a certain point in proficiency, at which the bilingual may access conceptual meaning directly. If the reader uses translation (L2 via L1) to process the sentence, the phonological and orthographic overlap would "lose potency", in the sense that the Norwegian word would be primed after the English one. The effect may also be explained through the IC model (Green, 1998), where the degree of inhibition in lexical access is influenced by the activation level of the lexical concepts in each language. As such, a higher level of activation causes a higher level of inhibition, where proficient L2 users display more inhibition than less proficient users.

The positive correlation between the inhibitory effect of priming in the begin-overlap conditions and the factor of *text proficiency* in the short conditions strongly agrees with the findings of Andrews & Lo (2012), in which high scores in text proficiency were associated with faster average reaction times, stronger inhibition from similar word primes and greater overall inhibition for transposed-letter words relative to neighbor primes. Especially spelling proficiency (which is included in our *text proficiency* factor) was associated with a stronger inhibitory effect from word priming. Andrews & Lo (2012) associates this with the concept of "Lexical Quality", which defines high quality lexical representations by principles of precision and redundancy (Perfetti, 1992). Precise representations are fully specified, so that the visual pattern of a word activates its lexical representation with minimal competition from similar words. Precision is enhanced by combining the orthographic and phonological words to form "stable" representations that define a word's identity (Andrews & Lo, 2012). Less

skilled spellers, however, showed facilitation for orthographically similar words (Andrews & Lo, 2012) – which is a finding that agrees with that of our own experiment.

For the factor of *L1 proficiency*, which includes spelling, grammar, reading writing and understanding skills, age of acquisition and exposure time, the positive correlation with begin-overlap conditions is also consistent with previous research (e.g. Van Heuven et al., 1998). According to the BIA model, lexical competition would only be resolved at the very last grapheme/phoneme of words with begin-overlap, which suggests that processing would be slower than if the word was unrelated in form. Language users who are highly proficient in L1 may transfer these skills (to a certain extent) to their use of L2. (For example, a person that considers spelling in their L1 to be important would likely be interested in spelling correctly in their L2 as well.) As the inhibition effect increased in tandem with the participants' proficiency levels, this is further evidence that high proficiency bilinguals react to stimuli in a similar manner to native users of the target language. Similarly, the positive correlation between begin-overlap items and factor 5 *General language proficiency* is consistent with previous research (e.g. Andrews & Lo, 2012; Marian et al, 2007; Van Heuven, 1998). This factor included a measure of L2 age of acquisition and time of exposure, which has been proven highly influential in the level of proficiency in L2 (e.g. Andrews & Lo, 2012; Marian et al, 2007).

General discussion

Our findings replicate and extended the findings of Frisson et al. (2014), showing that bilingual reader show inhibitory effects of form overlap in their L2 English reading as L1 English readers and that these effects were modulated by aspects of language profile and proficiency.

Inhibition was found in several measures, and although not all measures reached the significance threshold of .05 in the t-tests, there were several effects that were borderline significant, such as the inhibition for End-Overlap items in the First Fixation measure of the Spill-Over region, or the facilitation found in the same region for Begin-Overlap items. The borderline significant effects found in the First Fixation, Regression Path and Total Time measures of the Short Lag conditions may be interpreted as indicators of trends, but further experiments (with a larger participant group) are necessary in order to further investigate these aspects of bilingual language processing. The small sample size, is indeed, the main weakness of this study as it means that the statistical power of the experiment was low, in the sense that the pool of participants of 30 should have been larger – preferably 60 or more (which we plan to collect). With a larger participant group, it is possible that some of the borderline significant effects will become significant (especially the near-significant measures in the First Fixation measures of the Spill-Over region in Short Lag conditions). Another issue is that most of the participants were University students. As such, the statistical data was provided by a group of what is (presumably) skilled readers compared to an average in Norway. Additionally, one might assume that a larger proportion of University students are proficient English users compared to the rest of the population. As such, our results only reflect the effects on Norwegian-English bilinguals in higher education – not on a general basis.

The effects of orthography or phonology by themselves are also largely unexplored in this experiment. Frisson et al (2014) examined this aspect of priming in their experiments and

concluded that both orthography and phonology must overlap in order to produce significant effects. However, as our study did not examine the effects of phonological and orthographic overlap separately, we cannot say that Norwegian-English bilinguals have the same requirement of orthographic-phonological overlap as monolinguals in order to cause an inhibition/facilitation effect. The BIA model (Van Heuven et al, 1998) suggests that we might find similar effects, however – processing at the feature and letter levels apply to phonemes as well, which means that processing is affected both at the orthographic and phonological level. There are also other types of overlap and word neighborhoods that might cause facilitation or inhibition effects in word processing, such as transposed letters or letter deletion (e.g. Andrews & Los, 2012; Frisson et al, 2014). Further experiments might yield interesting findings on the workings of lexical neighbors and form overlap in Norwegian-English bilinguals. Additionally, Frisson et al (2014) looked at the effects of priming across sentence boundaries, and found that the effect disappeared. One might assume that a similar effect might be found in Norwegian-English bilinguals, especially those highly proficient in their L2 (as suggested by the RHM, Sunderman & Kroll, 2006). There is also the case of other punctuation, such as commas, which has been proven to remove the effects of priming in monolinguals (Carroll & Slowiaczek, 1986). It is likely that the same effect would appear in Norwegian-English bilinguals, but further testing is required to establish this for certain.

The assessment of the bilingual profile in our experiment is based on the results from the LEAP-Questionnaire. There are, however, other sides to self-assessment of language proficiency than those attended to by the LEAP-Q. Anxiety, for example, has been shown to affect performance in a variety of tasks (e.g. MacIntyre & Gardner, 1994; MacIntyre, Noels, & Clément, 1997). Also, it is worth mentioning that our study did not include any objective measures of language proficiency to control the self-assessment measures. Although self-assessment has been shown to be a reliable source of language skill assessment (e.g. Andrews & Lo, 2012; Marian et al, 2007), having separate objective measures of proficiency might have contributed valuable information on the participants' bilingual profile.

There is also the question of whether national or regional culture affects self-assessment. "Janteloven", a norm in which "one should never be better than, different to, or consider oneself superior to others", is a cultural phenomenon in Norway known to affect the perception of others in social settings (Bromgard, Trafimow, & Linn, 2014). It is possible that this also applies to bilingual self-assessment questionnaires, and might have caused participants to assess their proficiency scores lower than expected. There might be a suggestion that "Janteloven" is present in our self-assessment data already – although all our participants were native Norwegian users, the mean self-assessed value of L1 proficiency, on a scale of 1 (very low proficiency) to 10 (very high proficiency), the mean value was 6,8 (see Handeland (2018) for the full table of mean values from the LEAP-Q). Although this factor should not have affected the relationship between these variables and performance, it should be taken into account when comparing proficiency rating to other groups of bilinguals. Another social factor that may have influenced ratings were a social desirability bias - if the participant wanted to be accepted by the examiner, they might have rated the language they have most in common with the examiner (in this case, Norwegian) higher than the other language (Delgado, Guerrero, Goggin, & Ellis, 1999). As we did not control for social desirability, however, we cannot say anything about it with certainty.

Conclusion

In conclusion, our data show that Norwegian-English bilinguals are affected by word form similarity in a manner similar to English monolinguals, and that high-proficiency bilinguals display a level of inhibition and/or facilitation similar to native users, proportionate to their L2 proficiency. We have also shown that word form priming, when overlapping in both the orthographic and phonological level, cause inhibition in Norwegian-English bilinguals. Lastly, we have shown that skilled L2 readers are negatively impacted by word overlap to a higher degree than less skilled L2 readers are.

References

Andrews, S., & Hersch, J. (2010). Lexical precision in skilled readers: Individual differences in masked neighbor priming. . *Journal of Experimental Psychology: General*, pp. 299-318.

Andrews, S., & Lo, S. (2012). Not All Skilled Readers Have Cracked the Code: Individual Differences in Masked Form Priming. *Journal of Experimental Psychology*, 152-163.

Bowers, J., Davis, C., & Hanley, D. (2005). Interfering neighbours: The impact of novel word learning on the identification of visually similar words. *Cognition*, pp. B45-B54.

Braun, M., Hutzler, F., Münte, T., Rotte, M., Dambacher, M., Richlan, F., & Jacobs, A. (2015). *The neural bases of the pseudohomophone effect: Phonological constraints on lexico-semantic access in reading.* Elsevier LTD.

Bromgard, G., Trafimow, D., & Linn, C. (2014). Janteloven and the Expression of Pride in Norway and the United States. *The Journal of Social Psychology*, pp. 375-378.

Campbell, L. (2013). *Historical Linguistics: An Introduction.* Edinburgh: Edinburgh University Press.

Carroll, P., & Slowiaczek, M. (1986). Constraints on semantic priming in reading: A reading time study. *Memory & Cognition*, pp. 509-522.

Castles, A., Davis, C., Cavalot, P., & Forster, K. (2007). Tracking the acquisition of orthographic skills in developing readers: Masked priming effects. *Journal of Experimental Child Psychology*, pp. 165-182.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). A dual route cascaded model of visual word recognition and reading aloud. . *Psychological review*, pp. 204-256.

Davis, C. (2001). The self-organising lexical acquisition and recognition (SOLAR) model of visual word recognition. . *Dissertation Abstracts International: Section B: The Sciences and Engineering*, p. 594.

Davis, C., & Lupker, S. (2006). Masked Inhibitory Priming in English: Evidence for Lexical Inhibition. *Journal of Experimental Psychology*, 668-687.

Delgado, P., Guerrero, G., Goggin, J., & Ellis, B. (1999, February). Self-Assessment of Linguistic Skills by Bilingual Hispanics. *Hispanic Journal of Behavioral Sciences*, pp. 31-46.

Dijkstra, T., & Van Heuven, W. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, pp. 175-197.

Dufour, S. (2008). Phonological priming in auditory word recognition: When both controlled and automatic processes are responsible for the effects. *Canadian Journal of Experimental Psychology*, pp. 33-41.

Encyclopædia Britannica. (2017, 07 25). *Germanic Languages.* Retrieved from Encyclopædia Britannica: https://www.britannica.com/topic/Germanic-languages#ref603774

Ferraro, F., & Lowell, K. (2010). *Boston Naming Test.* Major Reference Works.

Forster, K., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology*, 680-698.

Frisson, S., Koole, H., Hughes, L., Olson, A., & Wheeldon, L. (2014). *Competition between orthographically and phonologically similar words during sentence reading: Evidence from eye movements.* University of Birmingham.

Grainger, J. (2008). *Cracking the orthographic code: An introduction.* Routledge.

Grainger, J., & Holcomb, P. (2009). Neural constraints on a functional architecture for word recognition. In P. Corneliussen, P. Hansen, M. Kringelbach, & K. Pugh, *The neural basis of reading.* Oxford: Oxford University Press.

Grainger, J., & Jacobs, A. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological review*, pp. 518-565.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition 1*, pp. 67-81.

Grosjean, F. (2010). *Bilingual: Life and Reality.* Harvard University Press.

Gutman, A., & Avanzati, B. (2013). *Languagesgulper.* Retrieved from Languagesgulper: http://www.languagesgulper.com/eng/Norwegian.html

Handeland, B., & Zitong, B. (2018). *Master's Thesis.* Kristiansand: Universitetet i Agder.

Hernández-Campjoy, J., & Conde-Silvestre, J. (2012). *The handbook of historical sociolinguistics.* John Wiley & Sons, Incorporated.

Kristoffersen, G. (2000). *The Phonology of Norwegian.* Oxford University Press. Retrieved from http://www.jstor.org/stable/4420207

Kroll, J., & De Groot, A. (2005). *Handbook of Bilingualism.* Oxford: Oxford University Press.

Lahiri, A., Wetterlin, A., & Jönsson-Steiner, E. (2005). Sounds Definite-ly Clitic: Evidence from Scandinavian Tone. *Lingue e Linguaggio*.

Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory and Cognition*, pp. 533-550.

Libben, M., & Titone, D. (2009). Bilingual Lexical Access in Context: Evidence From Eye Movements During Reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, pp. 381-390.

MacIntyre, P., & Gardner, R. (1994, June). The Subtle Effects of Language Anxiety on Cognitive Processing in the Second Language. *Language Learning*, pp. 283-306.

MacIntyre, P., Noels, K., & Clément, R. (1997, June). Biases in Self-Ratings of Second Language Proficiency: The Role of Language Anxiety. *Language Learning*, pp. 265-287.

Marian, V., Blumenfeld, H., & Kaushanskaya, M. (2007). *The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals.* Evanston: American Speech-Language-Hearing Association.

McClelland, J., & Rumelhart, D. (1981). An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Accound of Basic Findings. *Psychological Review*.

Paterson, K., Liversedge, S., & Davis, C. (2009, February). Inhibitory neighbor priming effects in eye movements during reading. *Psychonomic Bulletin & Review*, pp. 43-50.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145-174). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Rice, C. (2005). *Norwegian stress and quantity: The implications of loanwords.* University of Tromsø.

Rogers, H. (2014). *The Sounds of Language: An Introduction to Phonetics (2nd ed.).* Routledge.

Schwartz, A., Kroll, J., & Diaz, M. (2007, March). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Language and Cognitive processes*, pp. 106-129.

Sunderman, & Kroll. (2006). *First Language Activation During Second Language Lexical Processing.*

Talamas, Kroll, & Dufour. (1999). Form related errors in second language learning: A preliminary stage in the acquisition of L2 vocabulary. *Bilingualism: Language and Cognition*, pp. 45–58.

Titone, D., & Connine, C. (1997). Syllabification strategies in spoken word processing: Evidence from phonological priming . *Psychol Res*, pp. 251-263.

Van Heuven, W., Dijkstra, T., & Grainger, J. (1998). *Orthographic Neighborhood Effects in Bilingual Word Recognition.* Academic Press.

Wetterlin, A. (2017, 09 20). EN-122 English Phonetics, 20.09.2017 Class 4 (RP phoneme inventory - comparison) [PowerPoint slides]. Kristiansand.

Wiederholt, J., & Blalock, G. (2000). *Gray Silent Reading Tests.* Ausin, Texas: PROED Inc.

Woodcock, R., & Muñoz-Sandoval, A. (1993). *Woodcock-Muñoz Language Survey Comprehensive.* Chicago: Riverside.

Wu, Y., & Thierry, G. (2010). Investigating Bilingual Processing: The Neglected Role of Language Processing Contexts. *Frontiers in Psychology*.

APPENDIX

Appendix I: Full list of Stimuli

Condition	Relatedness	Item	prime	control	stimuli	Question
1	1	1	fork	knife	William used a big fork to eat the pork they had bought.	Did William eat pork?
1	1	2	brown	blue	The scary house was brown with a golden crown painted over the door.	Was a crown painted on the door?
1	1	3	settle	decide	They would finally settle to test their mettle against each other.	
1	1	4	kettle	soup	Sarah brought a terribly hot kettle with some nettle and other herbs in it.	
1	1	5	wall	desk	We found the yellow wall at the local mall and started painting.	
1	1	6	jump	wait	It was a long jump after the big hump in the road.	Was the hump in the road small?
1	1	7	mile	month	There was about a mile left before the file would be delivered safely.	
1	1	8	pile	heap	The clothes were left in a pile while the vile person was taking a shower.	Was the person taking a bath?
1	1	9	nurse	nanny	As she left her shift the nurse began to curse the difficult children.	
1	1	10	game	song	James played the game and gained fame all over the world.	Did James become famous?
1	1	11	flight	scene	They were too late for their flight after taking a slight left instead of going straight ahead.	
1	1	12	porch	grave	They sat by the porch with the unlit torch and started to cry.	Was the torch lit?
1	1	13	growl	whine	Panthers tend to growl a bit as they prowl around the forest.	
1	1	14	point	show	Chad made a point by smoking the joint in a public place.	
1	1	15	switch	return	Susan wanted to switch as her eyes would twitch when she looked at digital screens.	
1	1	16	cruise	trip	They came back from the cruise without a bruise even though the sea was rough.	
1	0	17	town	tailor	The bride walked to the tailor with her new gown in a shopping bag.	
1	0	18	sold	sent	They went ahead and sent tickets to a bold man they met outside.	
1	0	19	bill	note	I got an extensive note for just one pill from the doctor.	
1	0	20	sake	aunt	It was for her aunt they tried to fake a message of love and affection.	
1	0	21	bridge	boat	We went to the boat and threw my fridge over the side.	
1	0	22	faint	crash	I was totally about to crash when I saw a saint walking on the water.	
1	0	23	block	shirt	Harry bought the expensive shirt and the old clock he saw yesterday.	Did Harry buy the old clock?
1	0	24	flush	sneeze	When I know someone hears me sneeze I tend to blush a little bit.	

1	0	25	pillow	stick	Jenny took her stick to the old willow and laid down to sleep.	
1	0	26	join	drink	Paul wanted to drink but needed a coin to pay for it in the end.	
1	0	27	date	rain	It seemed like the rain was determined by fate when they first met.	
1	0	28	tired	angry	Leonard was so angry that he fired his worst employee.	
1	0	29	carry	accept	There are many responsibilities to accept if you choose to marry at a young age.	
1	0	30	haze	smoke	They ran through the smoke and found a maze in the middle of the forest.	
1	0	31	handle	carriage	She quickly turned the carriage after the lit candle fell to the floor.	
1	0	32	beach	shore	He went to the shore and ate a peach and a home-made pie.	Did he eat an apple?
2	1	33	seven	eight	He counted up to seven knights that could sever the dragon´s head.	Did he count the number of knights?
2	1	34	sheep	goats	The people watched the sheep to see the sheer weight of the wool they had on them.	
2	1	35	beam	glow	They could see a bright beam coming from the beak of a parrot statue in the hall.	Was the statue of a peacock?
2	1	36	rose	lily	Clara raised the rose closer to her nose to smell it properly.	
2	1	37	barn	shed	There is always a lot of noise at the barn when the dogs bark at the sounds outside.	
2	1	38	coat	glove	John had ruined his coat because of the coal he picked up in the mine.	Did John pick up the coal?
2	1	39	sharp	heavy	Chris brought a really sharp knife to kill the shark that was circling the boat.	Did Chris want to kill a whale?
2	1	40	spin	blend	Whenever she sat down to spin wool she spit her tobacco into a bowl.	
2	1	41	bowl	dish	Jonathan ate from a bowl with two large bows hanging over his shoulder.	
2	1	42	green	nice	The assistant wore a green jacket to greet everyone on his first day.	
2	1	43	hook	frame	David attached a hook to hang up the hood which he wore every day.	
2	1	44	seat	chair	It was hard to get a seat to watch the seal that could jump nine hoops in a row.	
2	1	45	sleep	dance	Ian wanted to sleep but then his sleek haircut might be ruined.	Was it Ian who had a sleek haircut?
2	1	46	mark	gift	Michael received an impressive mark for explaining the mars expedition to the audience.	Did Michael explain the moon expedition?
2	1	47	bulk	area	While the scientist measured the bulk of the shape a bulb went out above them.	
2	1	48	speak	shoot	When it was his turn to speak a metal spear was thrown at him.	

2	0	49	turf	gang	The long ongoing gang wars began with a turd being tossed towards someone at the Zoo.	
2	0	50	main	tiny	If you drive down the tiny street you see mail all over the road after the accident.	
2	0	51	trap	beat	Alice suspected that they would beat her in the tram station when she got there.	
2	0	52	stream	road	They went down the road after a winning streak at the rowing club	
2	0	53	bean	cube	Katie found a strange cube next to the bead she had lost earlier today.	
2	0	54	wing	foot	The strange bird would shake one foot and sort of wink at him from the high branch.	Was the bird on a low branch?
2	0	55	warm	cool	Keep the air in your bedroom cool in order to ward off sleepless nights.	
2	0	56	rail	tree	Carl leaned on the tree and watched the raid that was taking place in the village.	
2	0	57	corn	peas	When we gathered the peas we found a cord buried in the ground.	
2	0	58	brook	lake	Jackie went down to the lake with a broom to try and sweep up the glass.	
2	0	59	step	gain	The new discovery was a huge gain in the field of stem cell research.	
2	0	60	moon	cloud	They looked at the cloud and their mood changed a lot.	
2	0	61	train	field	They had to find the right field to follow the trail all the way home.	Were they going home?
2	0	62	boot	roar	There was a big roar as his boot hit the concrete floor.	
2	0	63	floor	path	Everything was lying all over the path after the flood hit the area we live in.	
2	0	64	dead	shot	The musician would rather be shot than become deaf before finishing his composition.	
3	1	65	wave	tide	Unfortunately a big wave forced them back out from the secret and unexplored cave on the beach.	Was the cave unexplored?
3	1	66	lack	miss	This equipment seems to lack everything that I told them was important for my rack of fishing nets.	
3	1	67	festered	swelled	His anxiety festered as he thought back on how he was pestered in high school.	
3	1	68	share	claim	The group felt that they had to share supplies even though the others would probably stare at them in surprise.	
3	1	69	crawl	slide	Due to the heavy snow the gang had to crawl back to their place after unexpectedly losing the brawl they had just been in.	Had the gang been in a brawl?
3	1	70	slug	toad	It looked like a slug had crawled around for hours on top of the plug in the kitchen sink.	
3	1	71	smell	aroma	There was a horrible smell which might have been a side effect of the spell he had just cast.	

3	1	72	sting	hurry	Bees tend to sting to defend the hive with all their strength as they fling themselves to their own deaths.	
3	1	73	freak	dude	There was a crazy freak following me around until I heard the soft creak of the shop doors closing behind me.	
3	1	74	patch	stamp	The alarm had a red patch that I noticed far too late after opening the hatch to the cellar of the house.	Did I open the hatch to the attic?
3	1	75	sound	noise	There was an overwhelming sound coming from the post office where a scary hound had been tied up to a pole.	
3	1	76	catch	judge	The police told us that they could catch the notorious criminal who had left behind a match at the scene of the crime.	
3	1	77	timber	forest	Cutting a lot of timber throughout his long career had made Fred unusually limber compared to others at his age.	
3	1	78	button	odour	Upon noticing the button Sally decided to select the box containing the mutton instead of the alternative.	Did Sally select the mutton?
3	1	79	race	move	They would race the longest distance they had ever attempted and knew that pace was the key to finishing.	
3	1	80	power	resolve	Faith can give you a lot of power so you never ever have to be scared and cower in fear from anything.	
3	0	81	rage	fury	Alison could feel the fury building up as she knew that her hourly wage was far below what she deserved.	
3	0	82	site	spot	At the landing spot there is a young boy who will always fly his kite whenever there is a little breeze.	
3	0	83	duck	goat	The fox became scared of the goat after losing the fight and finally decided to tuck its tail between its legs and escape.	Did the fox win the fight?
3	0	84	fist	club	Without a second thought John threw his club and managed to hit his opponent on the nose despite the mist that surrounded them.	Did John miss his opponent?
3	0	85	page	plan	When they looked closer at the plan in the old book there was a description of an unpleasant cage from medieval times.	
3	0	86	brief	quick	Ian wanted to be quick because he noticed that some people had expressed grief due to a lack of direction.	
3	0	87	fool	jerk	Everyone thought he was a jerk for letting himself to be used as a tool by the crooked administrator.	
3	0	88	notion	theory	Most of the villagers accept the theory that there is no such thing as a magic potion to remedy bad luck.	
3	0	89	danger	threat	The group might face some threat even though they are in the company of a ranger from the area.	
3	0	90	running	driving	These questionable people are driving a large pyramid scheme that was set up by a cunning banker who got fired.	Is the pyramid scheme large?
3	0	91	stay	rest	They needed a place to rest and a calm environment where they could sit and pray without being disturbed.	

3	0	92	peak	rise	They noticed a rise in petty crime and tried to prohibit anyone to leak it to the local media.	
3	0	93	pluck	raise	The chicken that Jack wanted to raise was in a bad mood and started to cluck when he advanced on it.	Was the chicken in a good mood?
3	0	94	sight	value	They both felt as if their value had decreased significantly after the very long and tight match that they had played.	
3	0	95	call	phone	Louise decided to phone the guy she met yesterday because he was tall and quite good-looking.	
3	0	96	hero	king	Connor was always considered a king for keeping the number of casualties down at zero throughout the disaster.	
4	1	97	paint	juice	Drinking the old paint caused Leif to experience lasting and severe stomach pains and he was rushed to the hospital.	Was Leif rushed to hospital?
4	1	98	purse	crate	The old but valuable purse contained a special crystal that people believed could purge any impurities from one's body.	
4	1	99	cheek	wrist	When Lee proposed, he stroked Sue’s cheek as he slid the ring on and a huge cheer arose from their families.	
4	1	100	market	shop	Sarah searched the market just up the street from me for a new marker to use for her notes.	Did Sarah need a paper towel?
4	1	101	slap	blow	The knight received a slap from the rest of his friends for his attempt to slam into those two friendly dragons.	
4	1	102	sweet	crazy	Elsa said it was sweet to see the girl who works in the clothes shop sweep the floor on weekends.	Did the girl work in a book shop?
4	1	103	chest	shelf	Rick searched a chest in his old countryside manor for a book about chess which he needed to study.	
4	1	104	clan	guys	He wanted to join the clan but he could only do that if he was clad in the proper clothing.	
4	1	105	grip	hold	Charles had a tight grip of the handle of the shopping cart and a huge grin on his face when they entered the mall.	
4	1	106	guilt	anger	Adam was feeling a lot of guilt about the theft that was committed today at the guild because of the unlocked door.	
4	1	107	cheap	royal	Ed knew the item was a cheap copy but he could not afford it and had to cheat to make sure he won the bid.	Did Ed know the item was a fake?
4	1	108	grain	fruit	A good supply of grain for the harsh winter season was the holy grail for the farmer following the poor harvest.	
4	1	109	tower	guard	He walked back to the tower of the castle during his day out to get his towel as he had forgotten it again.	
4	1	110	sees	gets	The university sees the monumental importance of encouraging students to seek further information on their own.	
4	1	111	brain	teeth	Sue made scans of the brain in her well equipped laboratory while twisting the braid in her long hair.	Does Sue work in a laboratory?

4	1	112	steam	fence	Due to the steam the thief and his slow and incompetent accomplice could steal the car without being noticed.	
4	0	113	sheet	cover	The morning’s thin cover of snow was quite dangerous but left a beautiful sheen on the new pavement that brightened the day.	
4	0	114	mount	steer	Behind the camel he was about to steer in order to cross the desert was a mound of big and smelly droppings he needed to avoid.	
4	0	115	paid	busy	He wanted to get busy but talking to the manager was a real pain every single time.	Did he like talking to his manager?
4	0	116	roof	cars	The earthquake shook the cars because it was so forceful and tossed the root of a tree through the window.	
4	0	117	queen	owner	Lee laughed out loud at the owner who had just arrived by car wearing a queer hat for the event.	
4	0	118	sign	book	They had looked at the book and decided together that they would never ever sigh at old people again.	
4	0	119	twin	mate	Martin was happy that his mate had gone into the woods and brought back a twig that was dry enough.	
4	0	120	leaf	flag	Clarence noticed a flag on the ground when he was about to leap over the edge of the cliff.	
4	0	121	team	star	The young basketball star did charity work and held back a tear of sadness for the poverty the people were living in.	
4	0	122	heat	cold	The wounded man suffered in the cold of the room while the poor woman tried to heal him as well as she could with few resources.	
4	0	123	hard	thin	The crystal she found was so thin that it could quite easily cause some serious harm to anyone who handled it carelessly.	Did she find a diamond?
4	0	124	word	piece	Brett couldn't read a specific piece in the old book he found in the attic because a worm had nibbled at some pages.	
4	0	125	lead	hire	If you are going to hire the crew you'd better make sure they know not to lean on the loose shelf behind the stage.	
4	0	126	hurt	sore	Several people got sore yesterday when an angry monkey suddenly decided to hurl rocks at the audience.	Did anyone get hurt?
4	0	127	curb	edge	She was sitting on the edge when someone came running by and cut off a curl of her hair all of a sudden.	
4	0	128	meat	flour	Daniel had become allergic to flour and had to be very careful about every meal he ate from now on.	
Condition	Relatedness	Item	prime	control	stimuli	Question
1	0	1	fork	knife	William used a big knife to eat the pork they had bought.	Did William eat pork?
1	0	2	brown	blue	The scary house was blue with a golden crown painted over the door.	Was a crown painted on the door?

1	0	3	settle	decide	They would finally decide to test their mettle against each other.	
1	0	4	kettle	soup	Sarah brought a terribly hot soup with some nettle and other herbs in it.	
1	0	5	wall	desk	We found the yellow desk at the local mall and started painting.	
1	0	6	jump	wait	It was a long wait after the big hump in the road.	Was the hump in the road small?
1	0	7	mile	month	There was about a month left before the file would be delivered safely.	
1	0	8	pile	heap	The clothes were left in a heap while the vile person was taking a shower.	Was the person taking a bath?
1	0	9	nurse	nanny	As she left her shift the nanny began to curse the difficult children.	
1	0	10	game	song	James played the song and gained fame all over the world.	Did James become famous?
1	0	11	flight	scene	They were too late for their scene after taking a slight left instead of going straight ahead.	
1	0	12	porch	grave	They sat by the grave with the unlit torch and started to cry.	Was the torch lit?
1	0	13	growl	whine	Panthers tend to whine a bit as they prowl around the forest.	
1	0	14	point	show	Chad made a show by smoking the joint in a public place.	
1	0	15	switch	return	Susan wanted to return as her eyes would twitch when she looked at digital screens.	
1	0	16	cruise	trip	They came back from the trip without a bruise even though the sea was rough.	
1	1	17	town	tailor	The bride walked to the town with her new gown in a shopping bag.	
1	1	18	sold	sent	They went ahead and sold tickets to a bold man they met outside.	
1	1	19	bill	note	I got an extensive bill for just one pill from the doctor.	
1	1	20	sake	aunt	It was for her sake they tried to fake a message of love and affection.	
1	1	21	bridge	boat	We went to the bridge and threw my fridge over the side.	
1	1	22	faint	crash	I was totally about to faint when I saw a saint walking on the water.	
1	1	23	block	shirt	Harry bought the expensive block and the old clock he saw yesterday.	Did Harry buy the old clock?
1	1	24	flush	sneeze	When I know someone hears me flush I tend to blush a little bit.	

1	1	25	pillow	stick	Jenny took her pillow to the old willow and laid down to sleep.	
1	1	26	join	drink	Paul wanted to join but needed a coin to pay for it in the end.	
1	1	27	date	rain	It seemed like the date was determined by fate when they first met.	
1	1	28	tired	angry	Leonard was so tired that he fired his worst employee.	
1	1	29	carry	accept	There are many responsibilities to carry if you choose to marry at a young age.	
1	1	30	haze	smoke	They ran through the haze and found a maze in the middle of the forest.	
1	1	31	handle	carriage	She quickly turned the handle after the lit candle fell to the floor.	
1	1	32	beach	shore	He went to the beach and ate a peach and a home-made pie.	Did he eat an apple?
2	0	33	seven	eight	He counted up to eight knights that could sever the dragon´s head.	Did he count the number of knights?
2	0	34	sheep	goats	The people watched the goats to see the sheer weight of the wool they had on them.	
2	0	35	beam	glow	They could see a bright glow coming from the beak of a parrot statue in the hall.	Was the statue of a peacock?
2	0	36	rose	lily	Clara raised the lily closer to her nose to smell it properly.	
2	0	37	barn	shed	There is always a lot of noise at the shed when the dogs bark at the sounds outside.	
2	0	38	coat	glove	John had ruined his glove because of the coal he picked up in the mine.	Did John pick up the coal?
2	0	39	sharp	heavy	Chris brought a really heavy knife to kill the shark that was circling the boat.	Did Chris want to kill a whale?
2	0	40	spin	blend	Whenever she sat down to blend wool she spit her tobacco into a bowl.	
2	0	41	bowl	dish	Jonathan ate from a dish with two large bows hanging over his shoulder.	
2	0	42	green	nice	The assistant wore a nice jacket to greet everyone on his first day.	
2	0	43	hook	frame	David attached a frame to hang up the hood which he wore every day.	
2	0	44	seat	chair	It was hard to get a chair to watch the seal that could jump nine hoops in a row.	
2	0	45	sleep	dance	Ian wanted to dance but then his sleek haircut might be ruined.	Was it Ian who had a sleek haircut?
2	0	46	mark	gift	Michael received an impressive gift for explaining the mars expedition to the audience.	Did Michael explain the moon expedition?

2	0	47	bulk	area	While the scientist measured the area of the shape a bulb went out above them.	
2	0	48	speak	shoot	When it was his turn to shoot a metal spear was thrown at him.	
2	1	49	turf	gang	The long ongoing turf wars began with a turd being tossed towards someone at the Zoo.	
2	1	50	main	tiny	If you drive down the main street you see mail all over the road after the accident.	
2	1	51	trap	beat	Alice suspected that they would trap her in the tram station when she got there.	
2	1	52	stream	road	They went down the stream after a winning streak at the rowing club	
2	1	53	bean	cube	Katie found a strange bean next to the bead she had lost earlier today.	
2	1	54	wing	foot	The strange bird would shake one wing and sort of wink at him from the high branch.	Was the bird on a low branch?
2	1	55	warm	cool	Keep the air in your bedroom warm in order to ward off sleepless nights.	
2	1	56	rail	tree	Carl leaned on the rail and watched the raid that was taking place in the village.	
2	1	57	corn	peas	When we gathered the corn we found a cord buried in the ground.	
2	1	58	brook	lake	Jackie went down to the brook with a broom to try and sweep up the glass.	
2	1	59	step	gain	The new discovery was a huge step in the field of stem cell research.	
2	1	60	moon	cloud	They looked at the moon and their mood changed a lot.	
2	1	61	train	field	They had to find the right train to follow the trail all the way home.	Were they going home?
2	1	62	boot	roar	There was a big boom as his boot hit the concrete floor.	
2	1	63	floor	path	Everything was lying all over the floor after the flood hit the area we live in.	
2	1	64	dead	shot	The musician would rather be dead than become deaf before finishing his composition.	
3	0	65	wave	tide	Unfortunately a big tide forced them back out from the secret and unexplored cave on the beach.	Was the cave unexplored?
3	0	66	lack	miss	This equipment seems to miss everything that I told them was important for my rack of fishing nets.	
3	0	67	festered	swelled	His anxiety swelled as he thought back on how he was pestered in high school.	
3	0	68	share	claim	The group felt that they had to claim supplies even though the others would probably stare at them in surprise.	

3	0	69	crawl	slide	Due to the heavy snow the gang had to slide back to their place after unexpectedly losing the brawl they had just been in.	Had the gang been in a brawl?
3	0	70	slug	toad	It looked like a toad had crawled around for hours on top of the plug in the kitchen sink.	
3	0	71	smell	aroma	There was a horrible aroma which might have been a side effect of the spell he had just cast.	
3	0	72	sting	hurry	Bees tend to hurry to defend the hive with all their strength as they fling themselves to their own deaths.	
3	0	73	freak	dude	There was a crazy dude following me around until I heard the soft creak of the shop doors closing behind me.	
3	0	74	patch	stamp	The alarm had a red stamp that I noticed far too late after opening the hatch to the cellar of the house.	Did I open the hatch to the attic?
3	0	75	sound	noise	There was an overwhelming noise coming from the post office where a scary hound had been tied up to a pole.	
3	0	76	catch	judge	The police told us that they could judge the notorious criminal who had left behind a match at the scene of the crime.	
3	0	77	timber	forest	Cutting a lot of forest throughout his long career had made Fred unusually limber compared to others at his age.	
3	0	78	button	odour	Upon noticing the odour Sally decided to select the box containing the mutton instead of the alternative.	Did Sally select the mutton?
3	0	79	race	move	They would move the longest distance they had ever attempted and knew that pace was the key to finishing.	
3	0	80	power	resolve	Faith can give you a lot of resolve so you never ever have to be scared and cower in fear from anything.	
3	1	81	rage	fury	Alison could feel the rage building up as she knew that her hourly wage was far below what she deserved.	
3	1	82	site	spot	At the landing site there is a young boy who will always fly his kite whenever there is a little breeze.	
3	1	83	duck	goat	The fox became scared of the duck after losing the fight and finally decided to tuck its tail between its legs and escape.	Did the fox win the fight?
3	1	84	fist	club	Without a second thought John threw his fist and managed to hit his opponent on the nose despite the mist that surrounded them.	Did John miss his opponent?
3	1	85	page	plan	When they looked closer at the page in the old book there was a description of an unpleasant cage from medieval times.	
3	1	86	brief	quick	Ian wanted to be brief because he noticed that some people had expressed grief due to a lack of direction.	
3	1	87	fool	jerk	Everyone thought he was a fool for letting himself to be used as a tool by the crooked administrator.	
3	1	88	notion	theory	Most of the villagers accept the notion that there is no such thing as a magic potion to remedy bad luck.	

3	1	89	danger	threat	The group might face some danger even though they are in the company of a ranger from the area.	
3	1	90	running	driving	These questionable people are running a large pyramid scheme that was set up by a cunning banker who got fired.	Is the pyramid scheme large?
3	1	91	stay	rest	They needed a place to stay and a calm environment where they could sit and pray without being disturbed.	
3	1	92	peak	rise	They noticed a peak in petty crime and tried to prohibit anyone to leak it to the local media.	
3	1	93	pluck	raise	The chicken that Jack wanted to pluck was in a bad mood and started to cluck when he advanced on it.	Was the chicken in a good mood?
3	1	94	sight	value	They both felt as if their sight had decreased significantly after the very long and tight match that they had played.	
3	1	95	call	phone	Louise decided to call the guy she met yesterday because he was tall and quite good-looking.	
3	1	96	hero	king	Connor was always considered a hero for keeping the number of casualties down at zero throughout the disaster.	
4	0	97	paint	juice	Drinking the old juice caused Leif to experience lasting and severe stomach pains and he was rushed to the hospital.	Was Leif rushed to hospital?
4	0	98	purse	crate	The old but valuable crate contained a special crystal that people believed could purge any impurities from one's body.	
4	0	99	cheek	wrist	When Lee proposed he stroked Sue’s wrist as he slid the ring on and a huge cheer arose from their families.	
4	0	100	market	shop	Sarah searched the shop just up the street from me for a new marker to use for her notes.	Did Sarah need a paper towel?
4	0	101	slap	blow	The knight received a blow from the rest of his friends for his attempt to slam into those two friendly dragons.	
4	0	102	sweet	crazy	Elsa said it was crazy to see the girl who works in the clothes shop sweep the floor on weekends.	Did the girl work in a book shop?
4	0	103	chest	shelf	Rick searched a shelf in his old countryside manor for a book about chess which he needed to study.	
4	0	104	clan	guys	He wanted to join the guys but he could only do that if he was clad in the proper clothing.	
4	0	105	grip	hold	Charles had a tight hold of the handle of the shopping cart and a huge grin on his face when they entered the mall.	
4	0	106	guilt	anger	Adam was feeling a lot of anger about the theft that was committed today at the guild because of the unlocked door.	
4	0	107	cheap	royal	Ed knew the item was a royal copy but he could not afford it and had to cheat to make sure he won the bid.	Did Ed know the item was a fake?
4	0	108	grain	fruit	A good supply of fruit for the harsh winter season was the holy grail for the farmer following the poor harvest.	
4	0	109	tower	guard	He walked back to the guard of the castle during his day out	

					to get his towel as he had forgotten it again.	
4	0	110	sees	gets	The university gets the monumental importance of encouraging students to seek further information on their own.	
4	0	111	brain	teeth	Sue made scans of the teeth in her well equipped laboratory while twisting the braid in her long hair.	Does Sue work in a laboratory?
4	0	112	steam	fence	Due to the fence the thief and his slow and incompetent accomplice could steal the car without being noticed.	
4	1	113	sheet	cover	The morning's thin sheet of snow was quite dangerous but left a beautiful sheen on the new pavement that brightened the day.	
4	1	114	mount	steer	Behind the camel he was about to mount in order to cross the desert was a mound of big and smelly droppings he needed to avoid.	
4	1	115	paid	busy	He wanted to get paid but talking to the manager was a real pain every single time.	Did he like talking to his manager?
4	1	116	roof	cars	The earthquake shook the roof because it was so forceful and tossed the root of a tree through the window.	
4	1	117	queen	owner	Lee laughed out loud at the queen who had just arrived by car wearing a queer hat for the event.	
4	1	118	sign	book	They had looked at the sign and decided together that they would never ever sigh at old people again.	
4	1	119	twin	mate	Martin was happy that his twin had gone into the woods and brought back a twig that was dry enough.	
4	1	120	leaf	flag	Clarence noticed a leaf on the ground when he was about to leap over the edge of the cliff.	
4	1	121	team	star	The young basketball team did charity work and held back a tear of sadness for the poverty the people were living in.	
4	1	122	heat	cold	The wounded man suffered in the heat of the room while the poor woman tried to heal him as well as she could with few resources.	
4	1	123	hard	thin	The crystal she found was so hard that it could quite easily cause some serious harm to anyone who handled it carelessly.	Did she find a diamond?
4	1	124	word	piece	Brett couldn't read a specific word in the old book he found in the attic because a worm had nibbled at some pages.	
4	1	125	lead	hire	If you are going to lead the crew you'd better make sure they know not to lean on the loose shelf behind the stage.	
4	1	126	hurt	sore	Several people got hurt yesterday when an angry monkey suddenly decided to hurl rocks at the audience.	Did anyone get hurt?
4	1	127	curb	edge	She was sitting on the curb when someone came running by and cut off a curl of her hair all of a sudden.	
4	1	128	meat	flour	Daniel had become allergic to meat and had to be very careful about every meal he ate from now on.	

Appendix II: Instructions for participants

Instructions for sentence reading

General information

In this experiment, we are going to monitor how your eyes are moving while you are reading sentences in English.

We are interested in your normal reading behaviour, so there is no need to go faster than normal. It is important that you read to understand the sentences. Therefore, every now and then you will be asked a question to test if you understood the sentence you have just read.

- If you are wearing any eye makeup, please remove it.
- Make sure that your hair does not get in the way of the screen (use the plastic hair band if you need to).
- A small sticker will be placed on your forehead to help the tracking.

Task events

- Sentences will appear on the screen, which should be read silently, at a normal pace, and for comprehension (you should understand them).
- When you have read and understood the sentence, look at the grey circle in the lower right corner of the screen and press the spacebar.
- Some sentences will be followed by a question. Answer them by pressing "1" for YES or "3" for NO.
- Other sentences will be followed by the statement "NO QUESTION". Proceed by pressing "1" (answer "YES").
- At certain points during the experiment, the screen will say "PAUSE". At these points in the experiment, there will be a short break (for as long as you need). When you are ready, press the spacebar to continue.
- When you read the sentences, try not to blink. (You can blink if you have to, but try to minimize blinking as much as possible while reading a sentence). Feel free to blink when a question or "NO QUESTION" appears.
- After each question (or NO QUESTION), a marker will appear on the screen to check if the camera is tracking correctly. Please look at the dot until it disappears.

Preparation

- Before the experiment can begin, a calibration and validation must be performed. A number of fixation points will appear on the screen. Look at each point until it disappears. (DO NOT ANTICIPATE THE NEXT DOT.)
- While reading the sentences in the experiment, your chin should be placed on the head rest, and your forehead should be touching the metal bar. Adjust the chair to make sure you are sitting comfortably in this position.
- Try not to move too much during the experiment. (If you need to move, blink, etc., please do it during the question segment).
- Place your index finger on "1" on the numeric pad of the keyboard and your middle or ring finger on "3". Your left hand should be on the spacebar.
- The experiment will begin with some practice sentences, followed by 4 sets of sentences. You can have a short break in between each set, after which a recalibration will be performed.

Appendix III: Instructions for experimenter

Eye tracking instructions

ALL SPEECH IS **ALWAYS** IN ENGLISH

Welcome the participant and ask them to leave their bags and jacket on a chair, and offer them a seat. Ask them to read the "Instructions for sentence reading" sheet, and inform them that "You don't need to remember everything, this is just to make you familiar with what we are going to do. I will repeat the most important bits when we get inside".

Go into the booth while the subject is reading. Turn on the desktop computer, and when it has started, turn on the laptop. (If the laptop doesn't continue into the tracker by itself, press "TRACKER".)

On the desktop computer, open the shortcut-folder "DAG latest ET version", and double click the file called "Exp1_EB_jan_asterix_long_deploy2". Enter the correct subject number, which follows this logic:

EXPERIMENT NUMBER, EXPERIMENT NAME, SUBJECT NUMBER, L, LIST NUMBER

It should look like this:

1ELL37L4

Then, select the correct list (the one you write in the subject number), and press ENTER.

Make sure that the head rest is placed in front of the camera. It should be fastened between the two tape strips on the desk.

On the tracker, press "Set options", "Select Config..." and choose "Desktop (Remote Mode), Target Sticker, Binoc/Monoc, 16/25mm lens, RBTABLER" (the fourth one from the top). Press "ACCEPT". Set the calibration type (top left corner) to 5-point and press "RETURN TO PREVIOUS SCREEN".

When this is done, return outside the booth and look moderately busy until you notice that the subject is done reading. Ask them to step into the booth and take a seat in the black chair. When they have, ask them to place their chin on the head rest, and to adjust the chair to make sure that they are sitting comfortably (the lever is on the left side).

When the subject is content with their seating position, use the up/down arrows on the laptop keyboard to adjust the sensor range, so that the dark blue only covers the subject's pupil. (If there are large dark blue areas outside the eye, the subject might be wearing makeup, which must be removed).

The light blue area shows the corneal reflection, and can be adjusted on the screen by using the arrows under "Corneal". Make sure that the white reflection in the eye has a light blue circle, but that nothing else does while the subject is looking in the vital areas (the top right corner will have some excess reflection, which is fine as long as the other areas are clean).

Stand next to the screen the participant is looking at, and tell the participant:

(While pointing where the sentence is going to appear): A sentence is going to appear on the screen. I'd like you to read it silently and for comprehension. When you've understood what the sentence says, there's going to be a big grey circle in the lower right corner of the screen (point to where the circle will be). When you've read the sentence and understood it, I want you to look at the grey circle, and while looking at it, press the SPACE BAR with your left hand. When you've done that, there's going to be a question here (point to where the question appears), or a statement saying NO QUESTION. If it's a question, I'd like you to press 1 (point to the number 1 on the numeric pad) to answer YES, or 3 (point to the number 3 on the numeric pad) for NO. If the screen says "NO QUESTION", I'd like you to answer yes, which is 1 (point to 1 again). When you've done that, a dot is going to appear in the middle of the screen. That dot is there to help me with the calibration, so I want you to look at it until it disappears."

Point to everything while explaining again now: "So just to repeat everything: A sentence appears in the middle of the screen. When you understand what it says, look at the grey circle, and WHILE LOOKING AT IT, press the SPACE BAR. When the question appears, you answer YES by pressing 1, and NO by pressing 2. Then, the calibration dot appears, and it's important that you look at it until it disappears."

"One last detail. When the sentence is in the middle of the screen, try your best not to blink (laugh a bit if they seem surprised). If you have to, you can blink. But try not to. When the question is on the screen, you can blink as much as you'd like."

"We're going to begin with 5 practice sentences, so you can get familiar with how this works. Then, we are going to do 4 sets of sentences. Between each set, the screen is going to say PAUSE. This means you can have a break for as long as you need."

Tell the participant that you are going to do a calibration, press the button on the screen that says CALIBRATE, and ask them to look at the dots until they disappear. When their eye focus has settled (about 1 second), press the SPACE BAR to accept the fixation. Repeat this for each step of the calibration, and press ACCEPT.

Tell the participant that we need to verify the calibration, and press VALIDATE. Repeat the process from the calibration. If the average value is above 0.50, you need to recalibrate and revalidate.

When the calibration is complete, tell the subject that we are going to begin, and press EXIT SETUP. This will start the test.

From now on, there are three things the experimenter needs to do:

1. If the participant stops, instruct them on how to continue
2. Press the SPACE BAR every time it pops up on the screen. Only press it if the participant is looking at it/close to it (or it's going to make a noise that will startle the participant).
3. If the light blue dot (the gaze pointer) seems to be outside the text or disappears, do a recalibration or adjust the CORNEAL reflection.

Make sure to compliment the participant on their performance when their screen says "This is the end of the experiment". Thank them for their help, and say that this is a really good data set.

If this is the last experiment the participant needs to complete them all, ask them to sign in the folder (next to their name), and while they do, fetch their gift card envelope. Thank them again for the help as you give it to them.

As they are leaving, ask them "By the way, would you be interested in being asked to participate in any other experiments at a later time?" If they answer yes, write "FEOK" (Further Experiments OK) next to their name in the folder. Thank them again and wish them a nice day.

Factor 1 (RC1): L2 text proficiency	*Loading values*	**Factor 2 (RC7): L1 proficiency**	*Loading values*	**Factor 3 (RC4): L2 speech proficiency**	*Loading values*	**Factor 4 (RC2): L1 text proficiency**	*Loading values*
Q121: L2 Exposure, reading	0.81	Q86: L1 Exposure, friends	0.78	Q132: L2 Proficiency, pronunciation	0.82	Q93: L1 Proficiency, reading	0.8
Q114: L2 Learning, friends	0.77	Q124: L2 Exposure, music	0.60	Q125: L2 Proficiency, speaking	0.75	Q91: L1 Proficiency, speaking	0.75
Q120: L2 Exposure, friends	0.77	Q98: L1 Regional accent	0.59	Q130: L2 Proficiency, vocab	0.66	Q94: L1 Proficiency, writing	0.65
Q126: L2 Proficiency, understanding	0.65	Q81: L1 Learning, reading	0.52	Q119: L2 Exposure, family	0.48	Q97: L1 Proficiency, spelling	0.54
Q128: L2 Proficiency, writing	0.60	Q96: L1 Proficiency, vocab	0.50	Q126: L2 Proficiency understanding	0.39	Q92: L1 Proficiency, understanding	0.54
Q127: L2 Proficiency, reading	0.60	Q100: L1 Importance, dialect	0.47	Q11: Age	0.39	Q95: L1 Proficiency, grammar	0.54
Q115: L2 Learning, reading	0.48	Q97: L1 Proficiency, spelling	0.47	Q129: L2 Proficiency, grammar	0.38	Q78: L1 Months in School	0.5
Q131: L2 Proficiency, spelling	0.47	Q87: L1 Exposure, reading	0.46	Q96: L1 Proficiency, vocab	0.37	Q99: L1 Others identify region	0.47
Q129: L2 Proficiency, grammar	0.41	Q95: L1 Proficiency, gramma	0.42	Q110: L2 Months with family	0.33	Q103: L2 Age begin acquiring	0.46
Q125: L2 Proficiency, speaking	0.39	Q117: L2 Learning, TV	0.38	Q108: L2 Months in country	0.30	Q11: Age	0.37
Q69: L2 into L1 Intrusions	0.38	Q101: L1 Modify accent	0.36	Q100: L1 Importance, dialect	-0.37	Q105: L2 Age began reading	0.37
Q85: L1 Exposure, family	0.32	Q118: L2 Learning, music	0.36	Q85: L1 Exposure, family	-0.41	Q34: L2 Percent Time used	-0.37
Q101: Modify L1 accent	-0.45	Q133: Self perceived L1 accent	0.34	Q79: L1 Learning, Family	-0.42	Q110: L2 Months with family	-0.63
Q11: Age	-0.37	Q128: L2 Proficiency writing	0.31	Q134: L2 Frequency identified	-0.65		
		Q136: L2 Effort with accent	-0.57	Q133: L1 self-perceived accent	-0.66		
		Q15: Years of formal education	-0.51				
		Q122: L2 Exposure, school	-0.45				
		Q103: L2 Age begin acquiring	-0.35				
		Q135: L2 importance accent	-0.34				
		Q105: L2 Age begin reading	-0.34				
		Q134: L2 Frequency identified	-0.33				
Proportion Variance	0.09	Proportion Variance	0.09	Proportion Variance	0.09	Proportion Variance Cumulative	0.08
Cumulative Variance	0.09	Cumulative Variance	0.18	Cumulative Variance	0.26	Variance	0.34
Factor 5 (RC5): L1 and L2 proficiency	*Loading values*	**Factor 6 (RC3): L1 learning**	*Loading values*	**Factor 7 (RC6): L2 learning speech**	*Loading values*	**Factor 8 (RC8): L2 age of acquisition**	*Loading values*
Q116: L2 Learning, school	0.69	Q84: L1 Learning, music	0.83	Q70: L1 into L2 Intrusions	0.73	Q106: L2 Age fluent reading	0.82
Q79: L1 Learning, Family	0.68	Q90: L1 Exposure, music	0.80	Q69: L2 into L1 Intrusions	0.68	Q104: L2 Age fluent speaking	0.75
Q80: L1 Learning, Friends	0.62	Q89: L1 Exposure, TV	0.77	Q113: L2 Learning. Family	0.67	Q103: L2 Age begin acquiring	0.46
Q131: L2 Exposure, spelling	0.59	Q83: L1 Learning, TV	0.73	Q119: L2 Exposure. family	0.58	Q105: L2 Age begin reading	0.39
Q95: L1 Proficiency, grammar	0.58	Q87: L1 Exposure, reading	0.51	Q123: L2 Exposure. TV	0.47	Q100: L1 Importance, dialect	0.34
Q129: L2 Proficiency, grammar	0.53	Q82: L1 Learning, school	0.43	Q124: L2 Exposure. music	0.46	Q135: L2 importance accent	0.34
Q122: L2 Exposure, school	0.51	Q81: L1 Learning, reading	0.42	Q34: L2 Percent Time used	0.46	Q116: L2 Learning, school	0.33
Q112: L2 Months in School/Workplace	0.48	Q15: Years of formal education	0.42	Q135: L2 importance accent	0.44	Q129: L2 Proficiency grammar	-0.31
Q97: L1 Proficiency, spelling	0.44	Q115: L2 Learning, reading	0.36	Q118: L2 Learning. music	0.43	Q82: L1 Learning, school	-0.39
Q94 L1 Proficiency, writing	0.43	Q117: L2 Learning, TV	0.34	Q117: L2 Learning. TV	0.35	Q92: L1 Proficiency, understanding	-0.49
Q85: L1 Exposure, family	0.39	Q118: L2 Learning, music	0.34	Q105: L2 Age began reading	0.32		
Q81: L1 Learning, Reading	0.37	Q101: L1 Modify accent	-0.4	Q127: L2 Proficiency reading	0.32		
Q108: L2 Months in country	0.36						
Q96: L1 Proficiency, vocab	0.35						
Q128: L2 Proficiency, writing	0.31						
Q105: L2 Age began reading	0.31						
Q110: L2 Months with family	-0.4						
Proportion Variance	0.08	Proportion Variance	0.07	Proportion Variance	0.06	Proportion Variance	0.06
Cumulative Variance	0.42	Cumulative Variance	0.50	Cumulative Variance	0.56	Cumulative Variance	0.62