Bilingual Visual Memory

The effects of cognates and hesitations on memory for visual scenes in proficient Norwegian-English bilinguals

IRENE SLETHEI

Master

SUPERVISORS
Professor Linda R. Wheeldon
Professor Allison L. Wetterlin

University of Agder, 2020
Faculty of Humanities and Pedagogy
Department of Foreign Languages and Translation

Table of Contents

List of figures

List of tables

Acknowledgments

No one could have predicted this level of turbulence in the year of MA thesis writing. Writing alone in a dark one-room apartment proved challenging when all the libraries and universities closed, and social distancing became the norm. Corona was a big bump in the road, but good people supported me nevertheless and made this thesis possible.

First, I must thank my advisors, Linda Wheeldon and Allison Wetterlin, for their never-ending support both in person and through video-meetings. Thank you for giving me the chance to participate in lab work, and for giving me experience with designing and conducting a linguistic experiment. I also want to thank Agnieszka Konopka for working with us on the design of this study, my co-experimenter, Andrea Skintveit Holte, for excellent cooperation in the project and the rest of the people in the lab for your support throughout this process.

Thank you to my kjæreste, my friends and family for letting me cry when I needed to cry, and for motivating me when everything felt dark and hopeless. And finally, thank you to all of you who agreed to participate in this study: nothing would be possible without your contributions.

Abstract

This study was designed to test the effect of spoken second language English processing on visual scene memory in proficient Norwegian-English bilinguals. In the study phase, participants looked at visual scenes and listened to descriptions of them in English while their eye-movements were recorded. After a break, they completed a judgment task (the test phase) in which they had to select which of two similar scenes they had seen in the study phase. We asked whether English speakers rely on visual information rather than verbal information when memorising input and whether this varied with proficiency and ease of processing. In particular, we manipulated the cognate status of the words in the scenes and the effects of dysfluencies in the auditory sentences prior to critical stimuli. We mapped proficiency levels and self-rated proficiency through a bilingual profile questionnaire and four objective language tests before the main experiment started. Our results show longer gaze durations in cognate and noncognate conditions after dysfluencies and imply that highly proficient L2 speakers' accuracy is facilitated by dysfluencies before noncognates early in the input, but that there might be a retrieval lag which builds up with increased amounts of input. The findings could suggest that early processing in highly proficient bilinguals resembles that of monolinguals, but that later processing is delayed into similar patterns as lower-proficient bilinguals. No cognate facilitation was found for memory or gaze durations, and our results imply that highly proficient bilinguals behave similarly to monolinguals in high-constraint sentence processing and that they might adapt into a more language-selective mode with the continuous L2 input. However, the results are merely trends in the data, and more research is needed before any conclusions are drawn.

"Die Grenzen meiner Sprache bedeuten die Grenzen meiner Welt":

"The limits of my language mean the limits of my world."

- Ludwig Wittgenstein, 1922

1 Introduction

More and more research on language processing looks at bilinguals rather than monolinguals and focuses on how knowing two languages causes certain advantages and disadvantages to the language users. For example, research on cognate words (i.e. words from two languages with similar spellings/pronunciations and shared semantic meanings, see section 1.2) shows that cognates facilitate processing whereas noncognate words (i.e. words that do not share similar spellings/pronunciations and semantic meanings) do not. Research has also proposed that bilinguals are better at inhibition- and attention control than monolingual language users (see e.g. Baumgart & Billick, 2018, more detailed in section 1.5). Further, a recent study (see Sampaio & Konopka, 2013) suggests that bilinguals remember details better in visual word processing than monolinguals (where both groups remembered the gist equally). Research on hesitations suggest that memory for what follows a hesitation is heightened, proposedly by the attention trigger the hesitations cause for a listener (see Corley, MacGregor & Donaldson, 2007). However, Konopka (internal report), did not find hesitation effects in bilinguals. The current study investigates the possible effects of cognates and hesitations on visual memory in Norwegian-English bilinguals. The study is motivated by older and newer research suggesting that cognate status and hesitations could influence memory for visual scenes in bilinguals. Participants in this study completed an augmented version of the bilingual profiling questionnaire ((Marian, Blumenfeld & Kaushanskaya, 2007) known as the LEAP-Q, with subjective proficiency ratings and screening as the primary goal, four language tests to map proficiency levels objectively and the main experiment investigating cognate and hesitation effects on visual memory. The three parts of the study will be described in detail in the method section.

1.1 What is bilingualism?

The term ‘bilingual’ can be used in multiple ways. *The Concise Oxford Dictionary of Linguistics* defines the term ’bilingual’ as someone “[…] with a native or native-like control of two languages” (Matthews, 2014, online), however, in the field of psycholinguistics, bilinguals are most commonly defined as language users who are able to understand and communicate in two languages. Someone who has a native-like control of two languages will be referred to as a ‘balanced bilingual’. Having two languages inside one’s head means that for each concept (for instance the concept of a cat, the concept of an apple), there are usually at least two lexical items to choose from; one from each language. When bilinguals process speech in one of their languages, does the other language interfere with the target language? Different hypotheses have been proposed to try and explain the process of lexical access in bilinguals, and in the following sections, these theories regarding lexical retrieval and models that try and account for bilingual processing will be reviewed.

1.2 Selective or nonselective activation?

Language users have what is known as the mental lexicon, which stores information on morphology, orthography, syntactics, semantics and phonology. The mental lexicon is a dictionary that language users access in order to process language. Different models have been proposed on how bilinguals' two languages could be stored inside the mental lexicon. Are lexical and semantic representations from the two languages stored together or in different parts of the mental lexicon (see e.g. Sánchez-Casas, R. & García-Albea, J. E., 2005)?

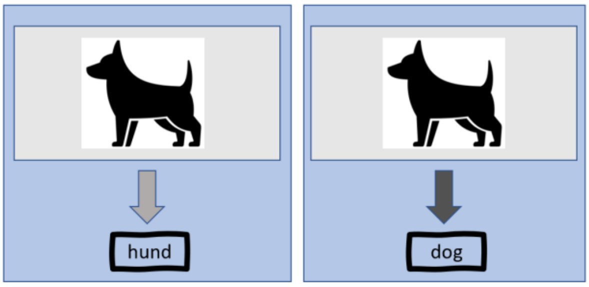

Figure 1, Lexical information and semantic information stored separately in the mental lexicon

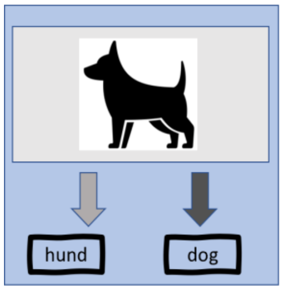

Figure 2, Lexical information and semantic information stored together in the mental lexicon

There are two main opposing hypotheses on language activation in bilingual processing. The languages might be exclusively activated and accessed separately (i.e. a Norwegian-English bilingual is unaffected by their second language (L2) when processing their first language (L1) and vice versa, see Figure 2), which is known as language-specific or selective language processing. This hypothesis supports a division of the two languages in the mental lexicon (ref. Figure 2) and suggests that the nonresponse language of a bilingual plays no role in processing.

The opposite view, which most research supports, is the language non-specific or nonselective processing hypothesis which proposes that the two languages have a more intimate connection and interact with each other during processing (i.e. a Norwegian-English bilingual is affected by their L2 when processing their L1) (see Dijkstra & van Heuven, 2002; Green, 1998; Kroll & De Groot, 2005; Costa et al., 2017). This supports, to some degree, the idea of shared concepts between languages in the mental lexicon (ref. Figure 2) and suggests that the nonresponse language of a bilingual might affect production and perception of the target language.

Research looking at nonselectivity has often used cognate status and interlingual homographs as their opposing stimuli in the search for evidence favouring nonselectivity in bilinguals (see e.g. DeGroot & Nas, 1991; Dijkstra, Van Jaarsveld & Ten Brinke, 1998; Dijkstra, Grainger & van Heuven, 1999; Costa et al., 2000;). Cognates are defined, in linguistics, as words that two languages from the same language family share with similar orthographic and/or phonological form and the same semantic meaning (e.g. the English word *cat* /cæt/ and the Norwegian word *katt* /kat/ both refer back to the same concept of a feline animal and originates from West Germanic) (from Oxford English Dictionary (oed.com), cat). However, in this study, cognates are used as is common practice in psycholinguistic studies, which excludes etymology as a criteria and only focuses on form and meaning, with the argument that we as language users do not know where the words originate from and will thus not use this information when processing language. Interlingual homographs, on the other hand, has similar to identical form between two languages but different semantic meanings (e.g. the English word *gift* (meaning something you give voluntarily to someone else), and the Norwegian word *gift* which translates to either 'married' or 'poison'). What research has found is a cognate facilitation effect and a homograph inhibitory effect (see e.g. Costa, 2000; Schröter & Schroeder, 2016), which is typically explained as evidence suggesting shared concepts in the mental lexicon where two cognates give double activation to their shared concept and interlingual homographs trigger both the target and nontarget languages to activate conflicting concepts which causes comprehension to take longer).

Libben & Titone (2009) investigated cognate facilitation and homograph inhibition in sentence context comprehension, where the motivation was to see if the effects often found in single-word processing could also be found in sentence processing, and if the effects would behave differently with either the biased or unbiased contextual information the sentences provided. Thirty highly proficient French-English undergraduate bilingual students from McGill University in Canada were tested in an eye-tracked reading experiment, where the goal was to look for nonselectivity in biased sentence contexts with cognate or noncognate manipulation. They also completed a language history questionnaire to map their level of proficiency to see if there might be a link between proficiency, facilitation and inhibition. The materials consisted of 32 French-English target cognates, 32 French-English target homographs and the respective 64 control words that were matched with their cognate or noncognate in terms of word frequency, length and neighbourhood density. The sentences had two clauses where the first clause was either biased or unbiased towards the target or control word that appeared in the second clause. Cognates and noncognates appeared in both a low-constraint sentence (i.e. sentence information was unbiased towards the target or control word) and a high-constraint sentence (i.e. sentence information was biased towards the target or control word), and the four sentence conditions (Cognate + Low-constraint, Cognate + High-constraint, Noncognate + Low-constraint, Noncognate + High-constraint) were divided into two lists so that each participant saw both the target word and its controlled partner but in different constrained sentences (i.e. when the target cognate was presented in a Low-constraint sentence, the control word for this cognate would appear in a High-constraint sentence for the same participant). To ensure that the participants read for content, a simple yes/no question appeared after 25% of the trials.

They found that, in the early stages of processing, both High-constraint and Low-constraint sentences showed cognate facilitation and homograph inhibition (around 350 ms after fixating on a word) whereas the selection process between 350-600 ms only showed facilitation and inhibition in Low-constraint sentences. They propose that language nonselectivity occurs early in lexical retrieval, but that the reader develops lexical expectations in High-constraint sentences which causes the process to evolve into a more selective mode due to the semantic restrictions applied to comprehension. In terms of proficiency, they found that the participants who were highly proficient in their L2 showed less cognate facilitation in High-constraint sentences, whereas no lesser homograph inhibition was found with increased proficiency. They propose that semantic constraints may promote language selectivity in the early stages of lexical access.

Lagrou, Hartsuiker & Duyck (2013) also looked at the possible influence of sentence context and the semantic constraints of a sentence on language processing, but in spoken word comprehension rather than orthographic word comprehension. In addition, as they had the option of manipulating the spoken word materials, they looked at how the native language of the speaker might affect comprehension. Their goal was to investigate three questions:

- Is there parallel activation of L1 and L2 when listening in one's L2?
- How do semantic constraints influence lexical access when listening to verbal sentences in one's L2?
- Are sub-phonemic cues provided by native accent from a speaker used by the listener to restrict lexical access when listening to sentences?

Sixty-four Dutch-English bilinguals from Ghent University took part in the study, whom all reported being quite proficient in their L2 although they lived in an L1-dominant environment. The target stimuli consisted of 30 Dutch-English interlingual homophones, 30 matched English control words, 60 English fillers and 120 nonwords following English phonotactic rules. Each target word appeared in both a high- and a low-constraint sentence context, and the stimuli were divided into two lists in order to ensure that participants only saw each target word once. The sentences were recorded both by a native Dutch speaker who were highly proficient in English and a native English speaker who was highly proficient in Dutch. The participants sat in front of a screen and saw a fixation cross before the audio played. After each trial, they were asked to decide whether the last word of the sentence was an existing English word or a nonword using one of two buttons in front of them. After they were done, they also filled in a questionnaire where they rated their L1 and L2 proficiency in different tasks (e.g. understanding, reading, and so on).

Results suggested that reaction times were slower when the last word was an interlingual homophone than the respective control words, which indicate activation and interference from the nontarget language and nonselectivity. Further, high-constraint sentences were processed faster, and homophones had a smaller, yet still present, delaying effect on reaction times. Lastly, the processing was faster when the sentences were pronounced by a native English speaker than by a native Dutch speaker, indicating that sub-phonemic cues closer to the listener's stored representations for their L2 aids word recognition. From these results, they conclude that lexical access is nonselective, and that both highly constraining sentence contexts and native accent compatible with the target language can cause fewer interactions between the two languages in processing (but not make it exclusively selective).

The two studies above suggest that the two languages of a bilingual interact to a bigger or lesser degree when presented in different sentence contexts visually or verbally. Level of selectivity seems not to be constant, but rather to vary in relation to the context in which the language is processed. Grosjean (1998) proposed the Language Mode hypothesis, which suggests that the level of nonselectivity depends on the bilinguals' language mode at the time of processing. The hypothesis assumes that a bilingual has two language modes: the monolingual mode (i.e. the mode where only one language is active) and the bilingual mode (i.e. both languages are active) (see e.g. Green & Abutalebi, 2013; Yu & Schwieter, 2018). For instance, if Norwegian-English bilinguals watch TV in English while sending a text message in Norwegian, both languages are utilised and thus active. On the other hand, if the same individuals are in a lecture with an English-speaking professor, write notes in English and read the powerpoint slides in English, L1 is not utilised and might receive less activation and affect L2 processing to a lesser degree as the English domination continues.

Language mode could alter the selectivity of lexical access, and bilinguals will thus score differently in tests and experiments if they have been in language switching contexts or L1/L2 exclusive contexts prior to, and during, the experiment. As cognates have shown to activate both languages, it is possible that cognates provoke a bilingual mode (as both languages activate) (see Dunn & Fox Tree, 2014). Dunn & Fox Tree (2014) looked at the plausible link between language dominance, language mode and processing. They found that bilinguals in a monolingual language mode rejected nonwords at the same rate as monolinguals, but that in the bilingual mode, they rejected them more slowly than both the monolinguals and the bilinguals in the monolingual mode. They also found that less proficient L2-users rejected the nonwords slower than higher-proficient L2-users.

1.3 Modelling bilingual lexical access and processing

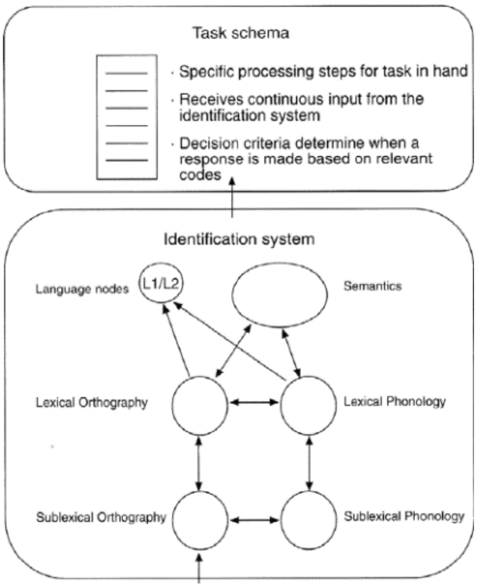

Figure 3, The BIA+ model as proposed by Dijkstra & van Heuven

Both Libben & Titone (2009) and Lagrou et al. (2013) found evidence that supports the BIA+ model proposed by Dijkstra & Van Heuven (2002); an upgraded version of their original BIA model (1998) which proposes that lexical access has both a word identification system and a task/decision system. Dijkstra & Van Heuven (2002) suggest that a word identification system activate words from both L1 and L2 based on orthographic, phonological and semantic representations. These parallel representations (the two language nodes in Figure 3) are linked to one another as well as to other relevant words before the orthography/phonology/semantics is recognised. When a bilingual knows two languages with the same orthographic writing systems (as e.g. Norwegian-English bilinguals), more words will activate from orthographic input than if the two languages do not share a writing system where the consequence might be that one language receives most of the activation. The BIA+ model is a cascading model which suggests that several words activate information on different levels and not only the target word. The task/decision system in the BIA+ model supervises the word identification system and inhibits inappropriate responses. The task/decision system also decides on the final, appropriate decoding of the input. Libben & Titone (2009) found that the initial lexical activation was nonselective, and that sentential context was fed into the task/decision system, giving comprehension a more selective form in high-constraint sentences. This is in line with the BIA+ model and the proposed task schema linked to semantic information which feeds back to an orthographic level. This in turn can exclude inappropriate responses based on semantic context in orthography. An issue with the BIA+ model as mentioned in Libben & Titone (2009) is that it does not account for the lack of cognate facilitation in high-constraint sentences. In theory, following the logic of the BIA+ model, cognate facilitation should still occur, as highly restricted sentence contexts do not exclude the cognate double link to concepts; cognates still have the same semantic meaning and would make sense in the context. They suggest that the BIA+ model might need feedback from the language nodes to the orthographic level as this could help explain the lack of cognate facilitation in high-constraint sentences.

The BIA+ model accounts for interaction between the two languages, but as seen above, proficiency affect cognate facilitation and interlingual homograph/homophone effects. The Revised Hierarchical Model (the RHM-model) of speech processing was proposed by Kroll & Stewart (1994) to account for the effects of proficiency on bilingual processing and the organisation and connections between the two languages in the mental lexicon. The RHM-model considers a nonselective and interactive approach, and proposes that both languages of a bilingual are linked to the same concept where L1 has a stronger bond to this common concept than L2, and that the bond from L2 to concept is strengthened with higher proficiency levels. The two languages are also connected, with the link from L1 to L2 being stronger than the link from L2 to L1. When a person becomes more proficient, the L1 and L2 become more balanced. A low-proficient bilingual depends more on their L1 to access concepts when processing their L2, whereas a highly-proficient bilingual can access concepts while processing their L2 with lesser or no help from their L1 (see Figure

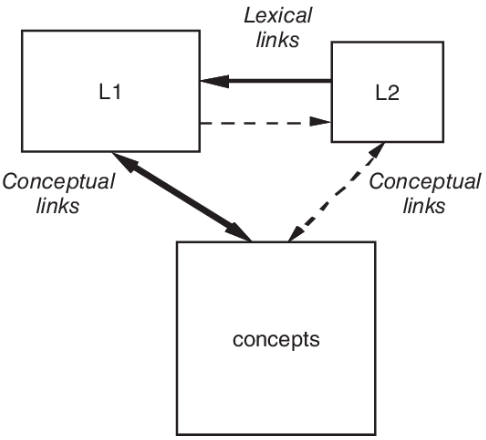

Figure 4, The RHM-model adapted from Kroll & Stewart (2004)

4). The RHM-model is different from the BIA+ model in their assumption that the lexical information from the two languages are stored separately with shared concepts, and Kroll et al. (2010) responded to criticism stating that the RHM-model is mainly a production model that does not exclude the BIA+ model but rather describes a different process. Nevertheless, the conclusions drawn by the RHM-model give a clear idea on how proficiency might affect processing in general.

Many studies have shown that different levels of L2 proficiency has an effect on bilingual processing (see e.g. Schröter & Schroeder, 2016; Sunderman & Kroll, 2006) and thus propose that the implications of the RHM-model are correct. The general findings show that when bilinguals become more balanced in their two languages, they depend less on their L1 knowledge when processing their L2. Language processing can be subdivided into two main categories: language production and language comprehension, and language comprehension has shown a less significant proficiency effect than production (see e.g. Kroll et al., 2002; Kroll et al., 2010; Schweiter & Sunderman, 2009). Language comprehension will be the main focus in this thesis.

It seems fitting to also briefly mention age of acquisition (often referred to as AoA) which is closely linked to proficiency levels. Research has shown that the age when a second language is acquired impacts proficiency and processing, and that later acquisition might contribute to weaker grammatical processing (see Hernandez & Li, 2007; van Hell & Tanner, 2012 for reviews). This means that controlling for AoA in experiments could be essential for the results to be as unanimous as necessary for reliability.

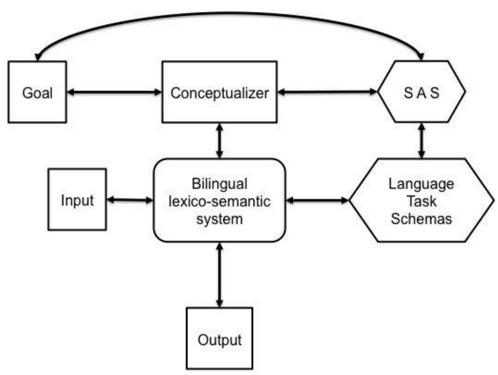

Figure 5, The Inhibitory-Control model, adapted from Green (1998)

The BIA+ model suggests how bilinguals access lexical items, the RHM- model hypothesise how the two languages are linked together dependent on proficiency levels, but how is it that bilinguals manage to suppress one language in order to only process in their target language? The Inhibitory-Control model (referred to as the IC-model) was proposed by Green (1998) and suggests that there are multiple levels of control in bilingual language processing. For Norwegian-English bilinguals to translate the word *'hund'* to the English word *'dog'*, they would have to inhibit the Norwegian version and control their focus and attention on the English equivalent. The IC-model hypothesises that every lemma (word base) is tagged with a language tag, i.e. *'hund'* is tagged with NORWEGIAN and *'dog'* is tagged with ENGLISH. The task schemas exclude all words with the wrong language tag for the task. Concepts activate two things: the lexico-semantic system and a supervisory attentional system (SAS), whose job is to control task schema activation in order to meet the goal of the processing. In other words, the SAS makes sure that the task schema activates the words (or lemmas) that are tagged for the Goal language (either L1 or L2) and inhibits the words with the wrong language tag. This inhibition and control creates competition between words that require attention from the language user, and the degree of activation from L1 and L2 lemmas decide the degree of Inhibitory-Control necessary for the task to be fulfilled. For instance, when Norwegian-English bilinguals process Norwegian, retrieving and processing the Norwegian word *'hund'* require less inhibition efforts than if they suddenly have to process the English word *'dog'* and need to inhibit their most dominant L1 already in focus. The three models mentioned above all try and account for different parts of the stages and organisation of processing, and by looking at their suggestions it is apparent that processing is more complex in bilinguals than if they were monolinguals with only one language to utilise. Bilinguals have more complex language processing than monolinguals which causes both advantages and disadvantages to the language user. Although the on possible advantages of being bilingual will be the primary focus, some disadvantages are shortly discussed before the positives are mentioned.

1.4 The bilingual disadvantage

Bilinguals generally have a smaller vocabulary in each of their languages than a monolingual has in their one language (Perani et al., 2003; Portocarrero, Burright & Donovick, 2007). In addition, lexical retrieval has been shown to be slower in bilinguals than monolinguals in picture naming tasks (see e.g. Kaushanskaya & Marian, 2007). Bilinguals also score lower on verbal category fluency tasks than monolinguals (see e.g. Portocarrero et al., 2007), which includes tasks where they are asked to name as many animals, foods, or other items within a category as they can. However, results are more mixed when it comes to letter fluency (i.e. name as many words as you know on the letter D) (see Sandoval, Gollan, Ferreira & Salmon, 2010). Bilinguals also experience more tip-of-the-tongue states (Gollan & Acenas, 2004), which is a state in which one has accessed the concept but not the lexical information connected to the concept (i.e. partial retrieval). TOT states in bilinguals are more present when processing in their nondominant language than their dominant language, but still they have more TOTs in their L1 than their monolingual peers (see e.g. Ecke, 2004; Gollan & Acenas, 2004).

There are some possible explanations to these disadvantages that relate to competition in selection and word frequency. The smaller vocabulary could be a result of frequency lag. Monolinguals build up word frequency in one language, whereas this frequency is shared between words in two languages in bilinguals. In other words, bilinguals use each of their two languages less than monolinguals use their one. This could cause both L1 and L2 processing to be slower as the frequency lag affects both languages (see Michael & Gollan, 2005; Gollan et al., 2011). Selection competition and smaller vocabulary could also interfere and make processing take longer than if the vocabulary was more extensive with only one word for each concept (e.g. Green, 1998; Kroll, Bobb & Wodniecka, 2006).

1.5 The bilingual advantage

On the positive note, there are many advantages to being bilingual, both cognitive and noncognitive. Being able to communicate in more than one language (and maybe especially in one of the 'Lingua Francas' like English and arguably Spanish and French) makes for opportunities when travelling, with political cooperations and applying for jobs (see e.g. report from New American Economy, 2017) amongst others. In terms of cognitive advantages, bilinguals have a higher metalinguistic awareness than monolinguals (see Adesope et al., 2010 for a review of relevant research; Campbell & Sais, 1995). In other words, bilinguals have a better ability to think about language, to be aware of different linguistic forms and to understand how language is composed to create meaning. In addition, research suggests that bilinguals also have a higher metacognitive awareness compared to monolinguals, which means that they have more insight and knowledge about their own cognitive processes in relation to learning strategies (Flavell, 1978; Kemp, 2007; Vorstman, De Swart, Ceginskas & van Den Bergh, 2009). This might be as a result of learning the different language aspects of an L2 (e.g. phonology, syntax, morphology) as well as how to use this information in context to communicate appropriately (Adescope et al., 2010).

However, the main advantages are linked to the higher levels of executive control found in bilinguals compared to monolinguals (Baumgart & Billick, 2017; Bialystok, 2011, 2009) in tasks that do not involve the use of language. Executive control can be defined as “the ability to carry out goal-directed behaviour using complex mental processes and cognitive abilities […]”(*Executive control* in Merriam-Webster dictionary, online), and it includes three core executive functions: inhibition control, working memory (Short-term memory) and cognitive flexibility (see Diamond, 2013). Inhibitory control involves the skills to control the attention and inhibit elements that might disturb the focus on the goal. E.g. a listener inhibits spoken words from one person to pay attention to another speaker. Three examples of tasks that show better executive control in bilinguals than monolinguals are the Stroop task, the Flanker task and the Simon task.

In a Stroop task, participants are requested to name the colour of a written word (which is the name of a colour), and the Stroop effect is seen through measuring the difference in reaction times between congruent (i.e. the written colour name and the font colour are the same) and incongruent (i.e. when the written colour name is not the same as the font colour) trials. The naming is done verbally. The Flanker task also measures the difference in reaction times between congruent and incongruent trials, but here the test subjects are asked to click on one of two keys in response to what direction the middle arrow out of five arrows is pointing. The congruent trials portrait five arrows, where the Flanker arrows point towards the same direction as the middle arrow (either (→→→→→) or (←←←←←)) and the incongruent trials consist of five arrows where one or more of the Flanker arrows point in a different direction from the middle arrow (e.g. (←←→←←), →←→←→)). In the Simon task, participants are asked to respond with either a left button or a right button to two conditions (e.g. press left when you see a left-pointed arrow and right button when you see a right-pointed arrow). The stimuli are then presented in congruent (e.g. the left-pointed arrow is on the left side of the screen) and incongruent (e.g. the left-pointed arrow is on the right side of the screen) trials, and the results are measured through the difference in response time between the two (congruent and incongruent trials). To be able to answer these tasks, people must inhibit the irrelevant information in the incongruent trials.

The Stroop task: Name the colours that pop up on the screen.
Correct answers: Red, Green, Blue.

Congruent trials:
Blue
Red
Green

Incongruent trials:
Red
Green
Blue

Bialystok, Craik & Luk (2008) conducted a study to assess working memory, lexical retrieval and executive control in 96 younger or older monolinguals or bilinguals. There were 24 younger bilinguals (mean age=19.7), 24 older bilinguals (mean age=68.3), 24 younger monolinguals (mean age=20.7) and 24 older monolinguals (mean age=67.2) participating in the study, and the tests on executive control involved the Simon arrows task and the Stroop task amongst others. The results of both tasks showed a difference in reaction times where bilinguals were faster in incongruent trials than monolinguals. The results of the Simon task suggested that the older monolinguals (61ms) struggled the most and that the older bilinguals (0.2ms) struggled the least. The old bilinguals thereby portrayed a higher amount of control in this task than the other three groups. The Stroop task showed less difference in reaction times between congruent and incongruent trials in the bilinguals than the monolinguals, and again, the older bilinguals seemed to experience lesser costs than their younger bilingual counter group.

Costa et al. (2009) conducted two experiments with the Flanker task on a group of 122 bilinguals and 122 monolinguals, where the first Flanker experiment had two test versions each with an overweight of one of the conditions (either 8 congruent and 88 incongruent trials per block, or vice versa). The second Flanker experiment also had two task versions, but here the first version had even numbers of the two conditions (48 congruent and 48 incongruent trials per block) and the second had 25% incongruent and 75% congruent trials (72 congruent and 24 incongruent). They found that in the low-monitoring conditions (i.e. conditions like in experiment 1 where one type of either congruent or incongruent trials dominated) there was no difference in bilinguals and monolinguals. However, in high-monitoring trials (i.e. when incongruent and congruent conditions shifted often throughout the blocks), bilinguals outperformed monolinguals. The conflict delay increased when there were more congruent than incongruent trials, and the bilinguals performed closer to the monolinguals as the conflict delay increased.

Bilinguals have better inhibitory control and attention control (Bialystok, 2009; Bialystok et al., 2008), which causes bilinguals to be better at problem-solving when the task involves inhibitory control, abstract thinking and sensory selection. This includes the ability to pay attention to relevant input in a noisy environment. Bialystok & Martin (2004) found that bilingual pupils had better attention control in a noisy classroom than their monolingual peers. They execute inhibition control and selective attention better than monolinguals, which could be explained by bilinguals' abilities to inhibit one language and move attention to the target language (the language with which they want to communicate).

Further, research has shown that bilingualism can help preserve executive control when aging, and help people diagnosed with Alzheimer's disease function at a higher cognitive level than monolinguals with the same disease (see Stern et al., 2005; Bialystok, 2011). Bilingualism may also cause Alzheimer's to develop a bit later in life than if a person was monolingual. A study by Bialystok et al. (2007) studied 184 people diagnosed with dementia and found that the bilinguals generally had shown signs of dementia 4 years later than the monolinguals (see Baumgart & Billick, 2018 for a review). The research mentioned provide evidence suggesting that bilingualism affects our brains, but could bilingualism also affect our memory for visual input?

1.6 Executive control and memory

Sampaio & Konopka (2013) recently found a link between bilinguals' executive control and memory. They investigated whether there is a difference in sentence memory in non-native speakers (bilingual L2-speakers) and native speakers (L1-speakers) of English as suggested by the RHM-model of language processing. They hypothesised that because L2-speakers are, to some degree, more dependent on the L2-L1 link in the mental lexicon than L1-speakers are the L1-L2 link, they encode the language more intensely and are thus able to remember surface form better than L1-speakers. Their hypothesis predicts that there should be no difference in the L2- and L1-speakers' ability to remember the gist of verbally presented sentences.

Their experiment was conducted on three groups of people: 26 monolingual speakers of English and 26 non-native speakers of English from the University of Illinois in the USA as well as 26 non-native speakers of English from Radboud Universiteit Nijmegen in the Netherlands. Their materials were in two sets of 12 + 12 sentences in either preferred form (i.e. the most commonly used synonym of a concept e.g. *hit*) or non-preferred form (i.e. the least common synonym of a concept e.g. *struck*):

Set A:

- 12 sentences in the preferred form
- 12 sentences in the non-preferred form

Set B:

- 12 sentences in the non-preferred form (the 12 first from set A, only changed the target word)
- 12 sentences in the non-preferred form (the 12 last from set A, only changed the target word)

Sentences from each set were divided into two 12-item lists, and the sentences within these lists were presented in the same order for each participant. In the test phase, participants received booklets with the start of every sentence in the presentation order to trigger their memory. They were presented with one 12-item list before told to write down the 12 sentences after their memory, and thereafter the next 12-item list.

Results showed that the non-native speakers recalled the correct surface form more times than the native speakers, and that the native speakers more often than the non-native speakers substituted the target non-preferred words with the expected synonym (the preferred word). There was no difference in recalling the sentences with the preferred synonyms between the native and non-native speakers. Their findings suggest that bilinguals processing in their L2 have better memory for details in sentence processing than monolinguals, however, because the native speakers of English were indeed monolinguals, we cannot tell from this study whether the effect in retention of surface form is an effect shown between monolinguals and bilinguals, or an effect shown between L2-speakers and L1-speakers. Further experiments are necessary where the L1-speakers are also bilingual to see if this effect is still measurable, or if it only occurs when testing monolingual L1-speakers. The RHM-model predicts that the effect should still be visible. The findings of Sampaio & Konopka (2013) imply that bilinguals remember verbal input more detailed than monolinguals, and the next question is how dysfluencies in verbal input might affect comprehension.

1.7 Memory and hesitations

Corley, MacGregor & Donaldson (2007) conducted a study investigating how hesitations in speech affect the listener and language comprehension. In daily communication, we produce and hear speech errors and hesitations, often in correlation with producing less predictable words, and Corley et al. (2007) hypothesise that listeners' memory is affected by dysfluencies in speech both short- and long-term. Hesitations might cause the listeners to be more attentive of the words following, as the pause often signals that the speaker is about to produce something more challenging than what has already been produced. The researchers further hypothesise that the N400-effect (i.e. the negative change in voltages in an ERP typically visible 400 ms after the onset of unpredictable words) should be reduced if the unpredictable word is preceded by a hesitation, thus making the unpredictable predictable. The study had 80 pairs of sentence frames with two final endings: One predictable and one unpredictable ending word. The sentences were then recorded both in a fluent and dysfluent (i.e. containing the filler *er*) way. The experiment consisted of two parts which were conducted on twelve native British English monolingual male speakers. In the first part of the experiment, the participants listened to a total of 160 utterances, and were instructed to listen for understanding. The recordings were presented in two blocks of approximately 15 minutes per block. After they were done, the participants were presented with the 160 utterances in writing with two alternatives for the ending word. The participants had to decide which word they had heard and which word was new by pressing one of two response buttons. Their accuracy was recorded.

The results showed that when the sentences had dysfluencies, the N400-effect was reduced, and they also showed a long-term effect in that the words following the dysfluencies were more often remembered than the words in fluent sentences. The latter suggests that the processing of fluent versus dysfluent sentences is different. However, this study does not examine the reasons as to why the N400-effect is lowered, and Corley et al. (2007) suggest two different explanations for their results:

- Post-lexical factors may be affected by the *er* causing the processing to happen after the word has been heard.
- The *er* may affect comprehension before the word is heard, reducing the extent to which specific predictions are made, and therefore increasing the integration difficulties.

The study strongly suggests that dysfluencies in speech affect the monolingual listener in language comprehension, both short- and long-term, however, in an internal report, A. Konopka reported that this was not replicated in bilinguals (study in print). When monolinguals hear dysfluencies, they remember the word following the dysfluency better, but Konopka's internal report suggests that bilinguals are not aided by dysfluencies in the same way. One possible reason for this is that bilinguals maximise their attention just by listening to L2 verbal input and thus have no capacity for more attention after a dysfluency. If this is the case, one should be able to find a proficiency effect. Lower-proficient L2 bilinguals might ignore the verbal input and just focus on remembering the visual input as the verbal becomes too much for them to process. Higher-proficient bilinguals should easier be able to process L2 verbal input, and they may or may not show a hesitation effect dependent on how much attention they are giving the processing of the L2 to begin with. Eye tracking can help show why or why not there is (or is no)t a hesitation effect in L2 processing of bilinguals. Theory on nonselectivity (especially with the cognate facilitation effect), executive control, bilingual visual memory and possible dysfluency effects in L2 processing motivate the study we designed and conducted to further investigate the effects of cognates and dysfluencies on visual memory and executive control. Our aim was to see if we could replicate the findings by Sampaio & Konopka (2013), Corley et al. (2007) and Konopka (in print) on a group of proficient Norwegian-English bilinguals.

1.8 Why Norwegian-English bilinguals?

Norwegian and English both originate from Indo-European and the Germanic Language family. This common origin from Indo-European and Proto-Germanic gives the two languages many cognate words from the basic word category such as numerals (e.g. *one, two, three* and *en, to, tre*) and astronomic objects (e.g. *sun, moon, star* and *sol, måne, stjerne*). The high occurrence of cognates is also a result of the Viking age from the start of the 8th century where English borrowed extensively from Old Norse (see Harbert, 2007). Words such as the English word *reindeer* (/ˈreɪndɪə/) comes from the Old Norse word *hreindyri* and is cognate with the Norwegian word *reinsdyr* (/ˈræinsdy:r/). Norwegian and English share several cognate pronouns as a result of the English borrowing, e.g. *that/det* (/ðæt/, /deː/) and *they/de* (/ðəɪ/, /dɛɪ/) where it must be considered that Norwegian has lost the dental fricatives /ð/ and /θ/. The verb *take* is borrowed from *taka* and exists in Norwegian as the verb *ta* (/ta:/). Because of this hugely shared vocabulary, it should be possible to observe a cognate facilitation effect, and possibly a better recollection of cognates than noncognates from memory. These two languages are, in other words, perfect languages in a bilingual to test for a cognate facilitation in memory as they share highly similar cognates. In addition, the languages share much grammar in terms of sentence building. Both languages are SVO-languages, although Norwegian is a strictly V2 language which needs the Verbal in second position in the structure. English accepts adverbials to be placed in the beginning of a structure without changing the order of the following clause elements, whereas Norwegian demands the Verbal to be placed straight after the adverbial.

- Lucy(S) read(V) a book(DO). (SVdO)
- Finally(A), Lucy(S) read(V) a book(DO). (ASVdO)
- Lucy(S) leste(V) ei bok(DO). (SVdO)
- Endelig(A) leste(V) Lucy(S) ei bok(DO). (AV2SdO)

1.9 The present study

Due to the many cognate words in Norwegian and English, Norwegian-English bilinguals can easily be tested for cognate facilitation, and in addition, the English proficiency in Norway is high, making it easy to find participants who master their L2 on a high level (see the EF English Proficiency Index, 2019). The bilingual profiles of the participants in our study are highly uniform as all Norwegians who attend the Norwegian educational system learn and acquire English from a very young age (five/six years old at the latest). Norwegian-English bilinguals who have completed primary and secondary school in Norway could be said to have a high level of proficiency in English, having learned the L2 in school for 11 years or more. The theory reviewed above predicts different processing results with different proficiency levels, and because our group of bilinguals have a high and uniform proficiency level, one can predict that they could show similar results to monolinguals. However, the proficiency levels are not identical which gives room for variation even in our group.

The first aim of our study was to further explore the effects (or lack of effects) of cognate status and dysfluencies in bilinguals L2 memory for visual scenes. We explicitly wanted to see if the memory of bilinguals was different from that of the monolinguals that were tested the sister project of Konopka (in print). Our second aim was to see if these effects (or lack of effects) would change with proficiency. All in all, our goal was to further find answers to two research questions:

- How does verbal input in terms of cognate status affect visual processing in bilinguals, and how might this correlate with proficiency levels?
- How does hesitations in speech affect bilingual processing, and how might this correlate with proficiency levels?

Research reviewed above make room for some predictions as to what the results will show, and the hypotheses and predictions based on this theory and the research questions are listed on the next page in three categories followed by a fourth null-hypothesis:

1. Cognate and hesitation effects on memory

- As a result of cognate facilitation in language processing, cognates should be remembered more accurately than noncognates.
- Hesitations will not affect bilinguals in processing in lower proficiency levels, but with increased proficiency levels, the bilingual listener might become more attentive of what follows, and thus show better memory for the object following.
- Eye movements will follow the audio throughout the scene, especially in cognate scenes and after hesitations.

2. Cognate and hesitation effects on gaze durations

- Cognates are easier to process and could by this show longer gaze durations than noncognates as they are recognised faster. This difference should decrease with lower proficiency.
- Hesitations might, with high proficiency levels, provoke longer gaze durations on the following object as seen in monolinguals (Corley et al., 2007).

3. Cognate effects on hesitations

- Hesitations should not have an effect in lower-proficient bilinguals but might have an effect with increased proficiency. If so, there should not be a clear difference in cognate and noncognate conditions as higher proficiency levels reduce cognate facilitation effects.

4. Null hypothesis

- Cognate status does not affect memory, gaze durations or hesitation effects.
- Hesitations do not affect memory or gaze durations.

2 Method

2.1 Participants

Thirty-three participants (of which 19 female and 14 male) with ages ranging from 18-32 (mean age=24.6) took part in this study, all of whom were Norwegian-English bilinguals with no other home languages than Norwegian (and possibly English if it was not the main home language). In order to be eligible for this study they had to confirm that they were between the ages of 18-35 and had normal or corrected to normal vision (including no colour-blindness) and hearing. Participants also confirmed to not have any language impediment diagnoses (e.g. dyslexia, stuttering), and to be reasonably good speakers of English. The participants were mainly students from the University of Agder, but family and friends of the experimenters were also included in the test group. The variation of participants should be seen as an advantage as the study does not only test university students, but also different, yet still highly proficient, bilinguals from other educational backgrounds. The testing was completed either in two days where the four language tests plus the LEAP-Q was finished the first day and the memory test was complete a different day, or all in one day where the participants could have breaks between the different tests when needed. Upon finishing all three parts of the experiment, the participants were reimbursed for their time with a 150 NOK gift card to be used in the University of Agder canteen.

2.2 LEAP-Q

2.2.1 Materials

The LEAP-Q is a general bilingual profile questionnaire developed by Marian et al. (2007) designed to assess the language profiles of bilinguals (and multilinguals) by collecting information regarding participants' background, language experience and attitudes towards their own language and language abilities. The LEAP-Q used in this study has been modified to also include a section about dialect and accent to collect data for a related PhD project in the language lab, but these data were not analysed for this study. In total, 14 questions were asked relating to screening, nine questions about language background, five about Norwegian English proficiency and 17 questions about dialect and accent. The LEAP-Q was filled in using Microsoft Office Exel (See Appendix B for the full adapted version of the questionnaire).

2.2.2 Design

The LEAP-Q consisted of four sections named 'screening', 'language background', 'Norwegian and English proficiency' and 'dialect and accent', which were the four areas of focus in the questionnaire. The screening section asked 14 general questions, such as what their age was, if they had normal or corrected-to-normal vision and hearing, and if Norwegian was their only home language aside from perhaps English. All these questions were asked to see if the participants met the criteria for participating in the study. The second section asked nine questions regarding their language backgrounds where they were asked to think about their language dominance in areas such as writing, spelling, dreaming and talking, and this section mapped how the participants were exposed to, or how they used, their two languages in their daily lives. They were asked how much percentage of the time they would choose to speak their two languages if they communicated with a bilingual who was equally as proficient as them in Norwegian and English. Section three consisted of five questions which focused on Norwegian and English only, and the participants had to answer how long they had been exposed to English and Norwegian in different environments and rate how proficient they thought of themselves in different language aspects of Norwegian and English (such as writing, spelling, reading, grammar, pronunciation etc.). Section four focused on dialect and accent of both Norwegian and English and had a total of 17 main questions which we did not use for our study.

The document was written in black letters on a white background, and the response boxes were light yellow before they turned white when the answers were filled in by the experimenter. Many of the questions were answered with a drop-down choice list (especially in the screening section), whilst others were typed in using the keyboard on the laptop. All of the responses were filled out by the experimenter. The main focal points for this study are language dominance, how much exposure the test group have of each language in different language usage situations and how they rate themselves in terms of proficiency in Norwegian and English.

2.2.3 Procedure

The LEAP-Q was conducted after the language tests described in the next section. The experimenter read the questions out loud and filled in the answers given by the participants. Some of the questions were more difficult to answer than others, and the experimenter then explained the questions in detail according to a pre-made protocol. The whole questionnaire took about 30 minutes. The subjects who completed all the experiments in one day were told to take a short break before the main test started. The others arranged a time to come back on a different day, and the time between the two testing days was usually between 2 days and two weeks from the first day of testing to make sure that their language proficiency did not undergo any changes between the pre-tests and the main experiment.

Table 1. Approximate durations of the different parts of the experiment

Test	*Approximate duration*
Norwegian vocabulary test	10 min
English vocabulary test	10 min
Working memory test	7 min
Lextale	3 min
LEAP-Q	30 min
Memory test	60 min

2.3 Proficiency tests

The participants completed all of the four proficiency tests in one sitting, but they could take breaks after every test if they needed to. These tests were designed to provide an objective measure of the participants' lexical proficiency in both Norwegian, English and their working memory. I will first describe the materials, design and procedure for each separate test before describing the overall procedure for conducting all four of the short tests.

2.3.1 Apparatus

All of the language tests were programmed using the experimental software package OpenSesame, version 3.1.9 (Mathôt, Schreij, & Theeuwes, 2012). The tests were done on two Lenovo ThinkPad T440 laptops provided by the University Experimental Linguistics lab. For the audio, the experimenters both used Sennheiser momentum M2 headphones.

2.3.2 Norwegian vocabulary test

2.3.2.1 Materials

This first task had two sections which each consisted of 20 Norwegian low-frequency noncognate target words (i.e. 40 different words made up the stimuli, see Appendix D2 for full word sets). Section 1 asked the participants to identify near-synonyms to the 20 target words whereas section 2 asked them to identify near-antonyms to the other 20 target words. Each trial in the two sections included one of the target words and five response options whereof four were single words and the fifth was an ‘I don’t know’ option. Only one of the words was the correct response, and the three other foils following each target word were either similar in meaning, its antonym in section 1 or its synonym in section 2.

2.3.2.2 Design and procedure

As mentioned above, the stimuli for this test were divided into two sections, where section 1 was always presented before section 2. The stimuli within each section were presented with a different randomisation for each participant, and there was a total of 20 trials per section. The target word in each trial was placed at the top of the screen, with the response options listed below with numbers 1-5 placed on their left side. The participants pressed the keys on the keyboard 1-5 depending on what alternative they chose, 5 always being ‘I don’t know’. The next trial started automatically when the participants pressed one of the response keys. This test only measured the accuracy of each participant and did not record reaction time. When section 1 was done, there appeared instructions on the screen for section 2, and the stimuli were presented as 24-pixel black text on a white background. When both section 1 and 2 were completed, the screen stated, ‘this task is done’.

2.3.3 English vocabulary test

2.3.3.1 Materials

The materials for this test were similar to the Norwegian vocabulary test materials, only now the target words and response options were English words (see Appendix D1 for full word sets). It had the same structure of two sections, whereof the first asked for near-synonyms and the second asked for near-antonyms. Even though the materials were similar to those in the Norwegian vocabulary test, the target words in the English test were not translations of the Norwegian target words, but rather 40 low frequency noncognate words unrelated to the Norwegian word stimuli.

2.3.3.2 Design and procedure

The design of this test was the same as the design of the Norwegian vocabulary test. It was administered after the Norwegian vocabulary test, and the instructions came up on the screen after pressing ENTER from the end screen of the first test. Section 1 was always presented before section 2, and the stimuli within each section were again presented with a different randomisation for each participant.

2.3.4 Auditory Working Memory task

2.3.4.1 Materials

In this test, two sequences of nonsense syllables were played for the participants, ranging from 5-7 syllables in length per sequence, and the task was to decide if the syllables in the two sequences occurred in the same or a different order. 144 nonsense syllables were made, and the syllables were constructed by a variety of vowels, single consonants and consonant clusters in both onset and coda position, all of which were language appropriate. The syllables within each pair of sequences were controlled to have different vowels and as few consonant repetitions as possible (e.g. baarrch, teeg, nup, gook all have different vowels and consonants in the onsets and codas) (See Appendix D3 for full list of stimuli). The audio was recorded by a speaker of Standard British English, and the syllables followed the phonological rules for English.

2.3.4.2 Design and procedure

The Working Memory test followed the two vocabulary tests described above, and the stimuli were presented auditorily to the participants through headphones. The same 24-pixel black text on white background was used, and whilst the recordings played, the screen was all white. After the two sequences of a pair had played, two options were presented on the screen: Press 1 if you think the order of the syllables was the same, and 2 if you think the order was different. The sequence pairs were played in a different pseudorandomised order for each participant with two constraints to the randomisation: There could be no more than three ‘same’ or ‘different’ trials after another, and two trials with the syllables switching in the same location could not follow each other. Between the two sequences in a pair, there was a 750 ms pause, and the next trial started automatically after the participants had pressed one of the response alternatives (1 or 2). In this test, accuracy was the only thing recorded. After the last trial, the screen stated, ‘this task is done’.

2.3.5 LexTALE

2.3.5.1 Materials

LexTALE (the Lexical Test for Advanced Learners of English) is a free standardised test for medium to highly proficient English L2 bilinguals designed to test for vocabulary knowledge and proficiency. The test is completed in English and has been thoroughly tested for reliability and validity (Lemhöfer & Broersma, 2012) where results suggested that Lextale provided more accurate proficiency predictions than self-ratings (see www.lextale.com for more information. This test only consists of 60 trials (excluding three dummy words), of which 40 items are existing words and 20 nonwords. The task is to decide if the combination of letters form an existing English word or not (see Appendix D4 for a complete list of stimuli words). E.g. ‘hurricane and ‘lofty’ are existing words whereas ‘exprate’ and ‘crumper’ are nonwords.

2.3.5.2 Design and procedure

The same 24-pixel black letters on white background was used in this test. In each trial, a string of letters appeared on top of the screen and two response alternatives showed up below the target word/nonword. The participants pressed 1 if they thought the letters formed an existing word in English and 2 if they did not think the letters formed an existing word of English. The words were presented in the same order for every participant, and the next trial started automatically when participants pressed either 1 or 2 on the keyboard. When they had finished all of the trials, the screen displayed, ‘this task is done’. The instructions proposed on lextale.com on how to implement the test and what text to inform the test subjects with were used in our experiment.

2.3.6 Language tests: Overall procedure

The participants were tested individually in a quiet room. Before the tests started, they were given an information sheet and a consent form (see Appendix A) where they accepted that their anonymised data could be used in this study for research purposes, and also that the anonymised data could be uploaded to an open access database. Participants signed two identical consent forms; one for the experimenter’s records, and one for them to take home.

They were asked if they were comfortable, if they wanted something to drink or if there was anything else they needed. Participants were told that they would complete four short tests on the laptop in front of them, and that the tests altogether would take approximately 20 minutes. They were also informed that they would complete a questionnaire upon completion of the language tests. Because the first test was in Norwegian, communication was in Norwegian before and during the Norwegian vocabulary test.

Prior to each test, instructions appeared on the screen, and the participants were instructed to read the text and to let the experimenter rephrase and explain the instructions before they started the tests. When the participants had completed the Norwegian vocabulary test, they were told that the communication would continue in English for the rest of the testing. Instructions appeared both before section 1 and section 2 of the two vocabulary tests, and the only difference was the language the instructions were written in and the language used in oral communication. During the rephrasing and explaining of the instructions of both sections in the Norwegian and English vocabulary tests, the experimenter emphasised the importance of pressing ‘5’ if they did not know the word, and to not guess if they did not know. This was important so that they did not select the right answer based on luck, but rather based on knowledge. When the two vocabulary tests were done, they read the instructions for the third test; the Working Memory test.

In the WM test, they were told to put on headphones and shown how they could adjust the volume on the laptop if they found the sound to be too low or high. They adjusted the headphones to fit their heads after they had received instructions. When they were ready, the experimenter pressed ENTER to start the first trial before the participants completed the test by themselves. After the last trial, they took off their headphones before the final language test.

The instructions for LexTALE were quite long, and the experimenter took time to ensure that all of the instructions were understood before pressing ENTER to start the test. It was explicitly stated that although the test was not timed, we were looking for first impressions rather than considered responses. As described in the design section, the next trial came up whenever the participant pressed the response keys, and ‘this task is done’ came up after the last trial of the test. The experimenter then pressed ENTER again to reveal a ‘thank you’ on the screen before recommending the participant to have a break for a few minutes before they started on the LEAP-Q. The screen was hidden from the participants when the experimenter again pressed ENTER once to reveal the results, and then a second time to close and save the data.

2.4 Visual memory task

2.4.1 Apparatus

In this experiment we used the SR-research Eyelink 1000 Plus (SR-Research) to record the eye movements of each participant, and a chin rest to minimize head movements. The experiment was designed and compiled by A. Konopka with the Experiment Builder (SR-Research) software.

2.4.2 Materials

The experimental materials consisted of three linked stimuli sets:

1. A matched set of 240 cognate and noncognate concrete nouns
2. A set of 80 photoshopped visual scenes in which pictures of the cognate and non-cognate objects appeared
3. Recorded sentence descriptions of the visual scenes with and without dysfluencies

An example of a pair of stimuli is shown in Figure 6, and the three stimuli sets are described in turn below.

Figure 6. An example of a cognate version (*pear*, *aubergine* and *eggs*) and a noncognate version (*orange*, *cucumber* and *cheese*) of the same scene.

Concrete nouns

Each scene pair required three cognate and three noncognate words, and the target word list therefore comprised 240 cognate words and 240 noncognate words whereof 80 of each were the ending items (i.e. the last cognate or noncognate mentioned) in their respective scenes (see Appendix F1 for full word pair stimuli set). The target words that did not fill in the final item spot had multiple constraints for their selection (see present study, section 1.9). Cognate and non-cognate lists were matched on length in terms of number of syllables (mean cognates= 1.78, noncognates=1.84), phonemes (mean cognates=4.86, noncognates=5.1) and word frequency (mean cognates=3.8, noncognates= 3.77) taken from the word frequency database SUBLEX-UK (see Van Heuven et al., 2014. for more information on this database). The ending item was never a critical item as it could not be eye tracked. The lack of a following item makes so that the eyes are not led to a new focus point, and thus the data would not be usable. The lead-in item was always a noncognate, and the same noncognate was used in both scene pair versions (e.g. in Figure 6, the noncognate *measuring jug* is the first thing mentioned in both of the versions). This was necessary in order to give the participants the same start independent on what version they were exposed to in their trial. The results would then not be biased by the start of the scenes as both participants saw and heard the same item in their respective scene of the pair. Because we wanted to see if memory for cognate words was better than for noncognate words, it was important that the noncognate words had approximately the same word frequency and length as the cognate words in the experiment.

The words were assigned to cognate and non-cognate versions of each scene. Within a scene, the objects depicting the words had to look as similar as possible (ref. Appendix F1). This means that each of the three cognate words in the cognate version had to be similar to the three noncognate objects in the noncognate version. The visual complexity of the objects and the placement of them had to match as closely as possible, including the size, colour and how well the objects fit into the specific scene. In Figure 6 above, the three cognate words *pear*, *aubergine* and *eggs* had their respective noncognates *orange*, *zucchini* and *cheese*. *Pear/orange* were matched as a pair, *aubergine/zucchini* as a pair and *eggs/cheese* as the third pair. In other words, both the scenes and the words within the scenes were paired, such that each of the three cognate words in a cognate scene had a noncognate partner in the noncognate scene with the same visual complexity, size and placement.

Scenes

There were a total of 101 scenes, whereof 80 were target scenes and 21 were fillers used so that the test subjects would not anticipate or discern any patterns in the critical scenes. The scenes were provided by A. Konopka who also advised on this study, and the scenes were similar to the ones Konopka used in an in-print study recently conducted. The scenes showed for instance a kitchen counter, a bathroom, a bedroom, a garden or other restricted areas and

Figure 7. Mirrored versions of the cognate version and the noncognate version in figure 6.

rooms that people are familiar with. In each scene there were different objects (i.e. the vocabulary items, ref. Figure 6) e.g. in a bedroom there might be a pair of slippers, a pair of pyjamas and a belt, together with normal bedroom objects such as a bed, a lamp, a closet and so on. Four objects were named in each scene, and they were always placed so that they were mentioned in a right-to-left or a left-to-right direction. The objects were placed in the scenes using photoshop, and it was important that the scenes and the objects looked natural and not manipulated. In addition to the cognate and noncognate versions of each scene, a mirrored version of each scene was created to prevent participants from anticipating gaze direction (see Figure 7 above for the mirrored versions of the scene in Figure 6). The sound files were not mirrored, but the visual scene could be either left-to-right or right-to-left (ref. Figure 6 and 7).

The fillers were more random scenes which only consisted of one recording with some similarities to the target recordings. Cognate status and word frequencies were not considered in the filler scenes(see Appendix F2 for full filler sentences). Some fillers had hesitations spliced into them, and these were added in much earlier or later than the hesitations in the target scenes (i.e. before the first object or the last object mentioned) to distract the participant from anticipating hesitation patterns in the target scenes. Some of the fillers had a different number of items mentioned than the target scenes, or focused on movements rather than concrete nouns (e.g. This is a hillside, there is a boy running and looking up at a kite).

As mentioned above, there was both a study phase and a test phase, and every target

Figure 8. Test phase versions of the two cognate/noncognate scene versions in figure 6.

scene had one main version and one changed version which was used in the test phase (see Figure 8 for the two test phase versions of Figure 6).

The changed version had one slight object change, and this could be for example that a dog was in a different position, that a sofa had changed (to a similar sofa), that a pile of sand was a bit darker in colour or that a poster had a similar but different motive. The differences were never major and were often hard to spot. Can you see the object change in the two scenes in Figure 8? The aubergine and the zucchini have small changes done to them. For each of the 160 (not counting the mirrored versions) scene versions, there were also 160 versions with one small (or sometimes bigger) change that was used in the test phase. The mirrored versions had the same change, only mirrored to fit the direction of the scene studied in the previous phase. The motivation for the changes was to make it difficult for the participants to spot the difference unless they had payed attention in the study phase. The changes were meant to be close to the original object, and only Object 2 and 3 were replaced.

Recordings

All the scenes had two recorded descriptions (one for the cognate and one for the noncognate version of the scene pair), and these were recorded by a native speaker of English with a mild Scottish accent. We used *Praat* (Boersma and Weenink 2019) to insert 500 ms between each mentioned object so that the participants had the same amount of time to process each object before the audio continued to the next object. The beginning of each sentence was identical in both the cognate and noncognate scene versions (i.e. the start of the sentence and the first noncognate mentioned were the same in both versions). In total, there were 181 different recordings; two recordings per target scene and one recording for each of the 21 filler scenes (see Appendix F2 for full sentence sets). The longest recording was approximately 13 seconds long. To construct the dysfluent descriptions, four different vocal hesitations were recorded (e.g. *eeeh*, *umm*) and one was spliced into each recording of a scene description 500 ms after the preceding object and 200ms before the next object. The same form of hesitation was spliced into the cognate and non-cognate version of each scene, and the hesitations were only spliced either before the second or the third object mentioned (i.e. only before the two critical objects), and always in the same position in each scene pair. This resulted in a total of four different recordings for each scene pair (see table 2 below): cognate and dysfluent, cognate and fluent, noncognate and dysfluent, and noncognate and fluent.

The first item mentioned was a noncognate before three cognate or three noncognate target words were mentioned. The audio was shaped as a sentence, and the following skeleton illustrates what the sentences included: "This is a bedroom/bathroom counter etc. There is a (noncognate filler word), a(n)/some (cognate1/noncognate1) +placement, a(n)/some (cognate2/noncognate2) +placement, and a(n)/some (cognate3/noncognate3) +placement". I will give two examples, where the first has the simplest structure (ref Figure 6, 7 and 8 for the pictures) and the second is more descriptive. Example 1: 'This is a kitchen counter. There is a measuring jug, a pear/an orange, an aubergine/a zucchini and some eggs/cheese'. Example 2: 'This is a street covered in snow. There is a large inflatable snowman, a tractor/car, a snowplough/truck, and a crow/seagull sitting on a railing'.

Table 2. The eight different versions of each scene

	Cognate status	Hesitation status	Hesitation status	Cognate status	
1	cognate	dysfluent	dysfluent	cognate	5
2	cognate	fluent	fluent	cognate	6
3	noncognate	dysfluent	dysfluent	noncognate	7
4	noncognate	fluent	fluent	noncognate	8

2.4.3 Design

As explained in the materials, each scene had a total of four items mentioned, and each scene pair had 7 items (with one shared noncognate and three cognates or three noncognates following). This experiment consisted of two parts: one study phase and one test phase. The study phase had a two-by-two design with cognate status (cognate/noncognate) and hesitation status (fluency/dysfluency) as crossed factors. In addition, all scenes could occur in mirrored form, therefore there were eight different versions for each scene (see Table 2 above).

The experiment had eight different lists (for both study and test phase), and each participant was assigned to one of them. Each list had 101 scenes comprising one version of each of the 80 target scenes and all the 21 fillers. The lists in the study phase showed the scenes in the same order where the first three and the final two scenes were fillers, and one filler appeared between each bundle of between 4-7 target scenes (see Appendix F2). Every participant was exposed to equal numbers of scenes in each condition. There were never two scenes with the exact same conditions following each other, however, two cognate scenes could follow each other if one was fluent and one dysfluent, or two dysfluent scenes could follow each other if one was cognate and the other one was noncognate and so on. All of the scenes appeared for 14 seconds. This was so that they could study the scenes for the same amount of time, so that the results could not be due to the differences in visual study time. To start a scene, the participants looked at a dot on the mid top of the screen and pressed space. Around midway there was a short pause where the participants could relax their eyes for a few seconds. The study phase included eye-tracking with an Eyelink (see apparatus, section 2.4.1), and participants' eye-movements were recorded during the 14 seconds the scene remained on the screen.

After the study phase was completed, the participants did 10 minutess of simple math problems before the test phase started. These problems involved simple subtraction and addition tasks (see Appendix G) and were added to prevent test phase being too simple for participants if they went straight to it from the study phase. In the test phase, the participants had to choose which image (see Figure 9 below) they had studied in the study phase. They were tested on each of the 80 target scenes, and the trials were presented with a different randomisation for each participant. One of the pictures on the screen was identical to the one that they studied whereas the other one had one small change applied to it.

2.4.4 Procedure

The participants were tested individually in a sound attenuated booth. They were seated in front of a stationary computer and the SR-research Eyelink 1000 plus, and the chair was adjusted so that they could sit comfortably while resting their head on a chinrest. The experimenter informed participants that they would have to sit in that position for about 35 minutes and that they would not be allowed to move their heads during the first part of testing (i.e. the study phase). The experimenter remained in the booth throughout, sitting behind the participant. Before the experiment, a white screen appeared, and the experimenter sat down in front of the control pc where the eye movements and scenes were visible to start calibration and validation of the eye tracker. They were told to look at different parts of the screen (e.g. top right, top left) so that the pupil sensitivity and the reflection of the camera could be adjusted appropriately. Then they were instructed to look straight at the black dot on the screen and follow it with their eyes, but to not guess where it would go next. They were also asked not to blink too much during the calibration.

When the eye tracker was calibrated and validated, the experimenter pressed the mouse and the instructions for the study phase appeared on the screen. They were told to read the instructions carefully before they were rephrased by the experimenter to ensure that the participants understood the task. The instructions said that they would now see a series of scenes and hear recordings for all of them, and that their task was to remember the pictures and the recordings to the best of their ability. They were told that they would see a total of 101 scenes which would take approximately 35 minutes, and that they would get a short break midway, where they should not move their head but rather relax their eyes for a few seconds. When they said they were ready, the experimenter pressed the mouse again and a black dot appeared on the top of the screen. They were instructed to look straight at the dot and press SPACE on the keyboard to start each scene. The recordings played on speakers placed next to the testing pc so both the experimenter and the participant could hear the recordings as the experiment ran. When the break came up, the participants were reminded to keep their head still. The camera was always recalibrated and revalidated before they continued after the break. Once they had studied all the pictures on the pc, the screen said to let the experimenter know they had reached the end of the first part of the experiment.

Immediately after the study phase, participants received a sheet of paper with simple maths questions (e.g. 52+5=, 7+14=, 167-4= etc. see Appendix G for full set) and were told to solve as many as they could in ten minutes. When ten minutes had passed, the experimenter collected the paper, and opened the test phase on the testing computer. Participants were informed that that the test phase would not be eye-tracked. The instructions were presented on the screen, and they explained how they would see two pictures of the same scene in each trial, and that their task was to indicate which of the alternatives they thought they had seen during the study phase. On the next screen there was an example of two scenes with only one object change, and text underneath telling them to press on of the following keys (see Figure 9 below):

- 1 (or 9) if they were sure they had studied the picture to the left (or to the right)
- 2 (or 8) if they thought they had studied the picture to the left (or to the right)
- 3 (or 7) if they guessed they had studied the picture to the left (or to the right)

Figure 9. Illustration of how the test phase screen looked like for the participants.

This text was visible during each of the trials as well. The participants were told to place their fingers on 1,2,3 and 7,8,9, and that no other keys would work during the test. When they were ready, the experimenter clicked on the mouse to show the first couplet, and the next trial showed up once the participant had pressed one of the answering keys. They ran the test themselves until the test screen thanked them for their participation and marked the end of the test phase and the experiment.

When testing was completed, the experimenter asked the participants if there were any aspects of the experiment they had reacted to, any words they did not know or any thoughts they had after participating, and the responses were written down so that we could use their feedback to see how well the experiment hid the manipulations, or if there were any weaknesses in the vocabulary items, design or procedure of the experiment.

3 Results

3.1 LEAP-Q

All participants ranked Norwegian as their most dominant language and the language that they first acquired. They also ranked English as their second most dominant language, although three participants reported acquiring a different language prior to English (of which two had acquired Danish and one Swedish). In terms of culture identification, all but one participant reported Norwegian as the culture they identified most with (mean: 8.6, range: 5-10). In terms of time spent in an English-speaking country, it ranged from 0 to 4 years. Table 3 shows the mean responses and ranges in the general screening, and includes task language, exposure and contributing factor ratings and ages for when different linguistic developments occurred.

Table 3. Means and ranges for self-rated language use and exposure

Question	Norwegian		English	
	Mean value	**Range**	**Mean value**	**Range**
Age	24.6	18-32		
Gender	19 F, 14 M			
Years of education	16.3	12.5-19		
Total exposure (% of the time)	61.5	30-80	36.7	19-70
Speaking (% of the time)	84.2	40-99	14.3	1-40
Reading (% of the time)	49.7	10-95	50.3	5-90
Language of choice (% of the time)	83.8	0-100	14.6	0-100
Maths language (out of 33 participants)	32		1	
Dream language (out of 33 participants)	31		2	
Anger language (out of 33 participants)	31		2	
Self-talking language (out of 33 participants)	30		3	
Exposure- family (from 0-10 of the time)	9.2	0-10	0.8	0-10
Exposure- friends (from 0-10 of the time)	8.3	5-10	2	0-6
Exposure- reading (from 0-10 of the time)	4.7	1-9	5.5	1-10
Exposure- TV (visual) (from 0-10 of the time)	3.2	0-7	7.1	3-10
Contributor- family (from 0-10)	9.4	3-10	2.5	0-10
Contributor- friends (from 0-10)	6.8	2-10	5.8	0-10
Contributor- reading (from 0-10)	6.9	3-10	7.8	4-10
Contributor- school (from 0-10)	7.9	3-10	7.5	2-10
Contributor- TV (visual) (from 0-10)	4.1	0-9	7.5	1-10
Contributor- music (audio) (from 0-10)	3.5	0-8	6.6	0-10

The mean responses for language of choice strongly favoured Norwegian, although the range was broad. Further, the vast majority reported doing maths, dreaming, expressing anger and talk to themselves mainly in their first language, Norwegian. The mean reported total exposure is also higher for Norwegian than English, but some participants rank their English exposure higher than the Norwegian exposure. Several of the participants were in the process of writing their English MA at the time of testing and would thus have a high exposure rate in English due to lectures, reading and writing mainly being in their L2. The exposure through friends and family is predominantly Norwegian, but the exposure through reading, visual and auditory media is reported as predominantly English. In Norway, much of films, series and music is in English. The answers to what contributed to their learning of each language reflects the answers regarding exposure. Family and friends are important contributors in the learning of L1 whereas reading, visual and auditory media are reported as being more important in the learning of L2 than L1. The contribution of school was similar for L1 and L2.

Table 4 shows the mean ratings and ranges on the questions that required them to rate their own proficiency levels in English and Norwegian, as well as how often they reckoned their two languages mixed when processing (both accidental and intentional).

Table 4. Means and ranges for self-rated language proficiency and language mixing

Vocabulary proficiency (self-rated from 1-10)	8.2	6-10	7.1	4-10
Spelling proficiency (self-rated from 1-10)	7.9	4-10	6.8	4-9
Speaking proficiency (self-rated from 1-10)	9.4	5-10	7.4	5-10
Pronunciation prof. (self-rated from 1-10)	9.2	6-10	6.6	3-9
Reading proficiency (self-rated from 1-10)	9.4	7-10	8.4	6-10
Writing proficiency (self-rated from 1-10)	8.3	5-10	7	4-10
Grammar proficiency (self-rated from 1-10)	7.9	4-10	6.7	4-9
L2 intrusion into L1 (accidental from 0-10)	3.3	0-8		
L1 intrusion into L2 (accidental from 0-10)			2.7	0-3
L2 inclusion into L1 (intentional from 0-10)	3.8	0-9		
L1 inclusion into L2 (intentional from 0-10)			2.1	0-5

The mean ratings in proficiency in different language tasks are lower in all of the English scores than the Norwegian scores. However, the scores in reading, writing, vocabulary and spelling are quite similar in the two languages. As can be seen from the range of responses, some of the Norwegian ratings are quite low for native speakers. Participants also reported more L2 mixing when producing L1 both accidentally and intentionally than L1 intrusions into L2 production.

3.2 Objective language proficiency tests

The results of the objective language proficiency tests are shown in Table 5. None of the participants scored 100% on any of the four tasks, and LexTALE is the task with the overall highest scores. The mean performance in Working Memory is relatively high, and the vocabulary scores are the most variable of the four, with a wide range of scores in both languages. However, the mean scores between the Norwegian and English vocabulary tests are similar with similar ranges. The lextale scores indicate that the participants have very good English vocabulary recognition.

Table 5. Means and ranges for the objective language proficiency tests

Test	Mean score	Range
Norwegian vocabulary test %	32.8	10.5-57.5
English vocabulary test %	31.7	7.5-57.5
Auditory WM task %	65.8	46.7-86.7
Lextale %	82.4	63.5-92.2

Feedback from the participants during and after the four language tests suggested that they found the Norwegian vocabulary test more challenging and difficult than the English vocabulary test. To determine the relationship between the participants' self-rated proficiency and their test results on the language tests, we ran correlations between their rating of their English vocabulary proficiency with test results from the English vocabulary test and the test results on LexTALE. The correlations are shown in Figures 10 and 11 respectively (next page). The English vocabulary scores and proficiency ratings showed a significant positive correlation, $r=0.43$, $p<.05$. The correlation between LexTALE and the English proficiency ratings showed an even higher positive correlation, $r=0.59$, $p<.0001$.

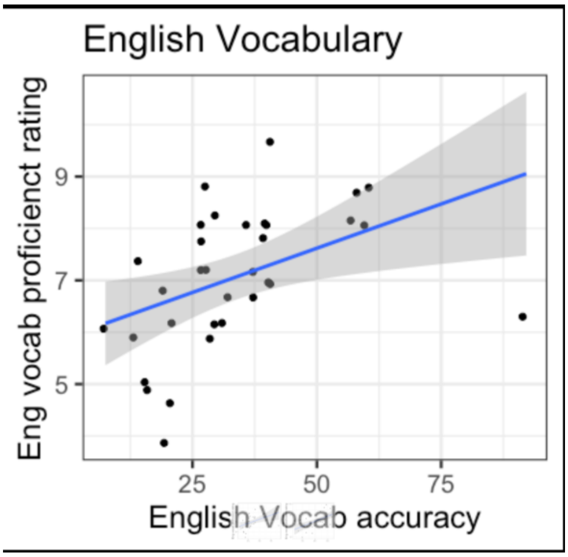

Figure 10, Correlations between L2 vocab accuracy and self-rated proficiency.

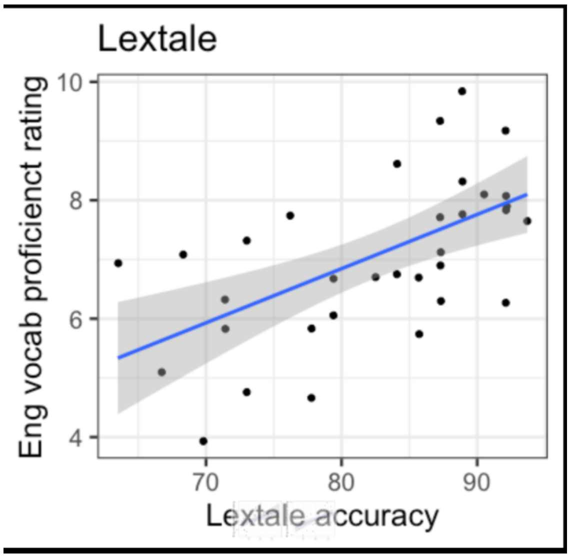

Figure 11, Correlations between LexTALE results and self-rated proficiency.

3.3 Visual memory test

3.3.1 Accuracy data from the test phase

The accuracy means for the means for the test phase are shown in Figure 12. The results are shown for Object 2 and Object 3 separately. As can be seen, the effect of cognate status on accuracy is neither consistent nor large (cognates 0.7975, noncognates 0.8025 overall). The main effect of dysfluency was also small (dysfluent 0.8025, fluent 0.7975 overall). The effect of dysfluency was different for Object 2 and Object 3. Accuracy was slightly higher in the dysfluent noncognate condition for Object 2. For Object 3, the same condition (dysfluent and noncognate) had a slightly lower accuracy than the noncognate fluent condition. A linear mixed effect model of the data was run including the experimental conditions as fixed effects and the objective English language tests as continuous factors. The best fitting model is shown in Table 6 on the next page. As

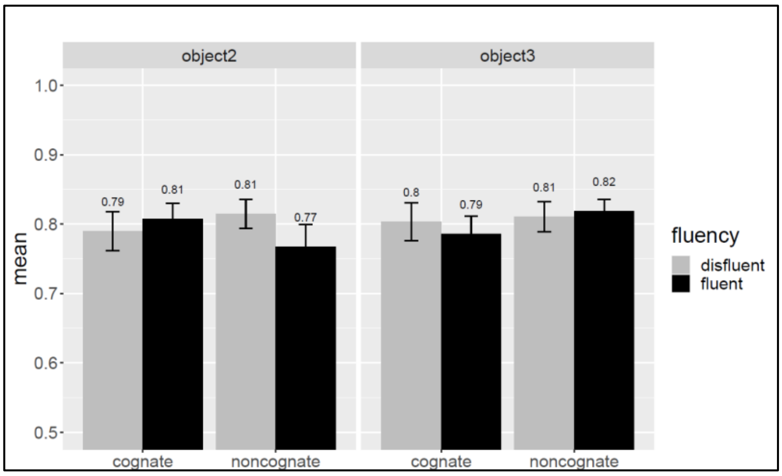

Figure 12. Mean and standard errors for the test phase shown object 2 and 3 for fluent and dysfluent trials for cognate and noncognate conditions.

can be seen, only the three-way interaction of Cognate status, Fluency and Object approached significance. None of the objective proficiency measures significantly contributed to the model.

Table 6. Best fitting model predicting the test phase accuracy data.

```
Random slopes for Cognate; models with other slopes do not converge.
Fixed effects:
                                 Estimate Std. Error z value Pr(>|z|)
(Intercept)                      -1.62569    0.12628 -12.873   <2e-16 ***
cognate1                         -0.01327    0.16335  -0.081   0.9352
fluency1                          0.07876    0.10497   0.750   0.4531
tested_object1                   -0.05932    0.18759  -0.316   0.7518
cognate1:fluency1                 0.11621    0.20990   0.554   0.5798
cognate1:tested_object1          -0.17916    0.27454  -0.653   0.5140
fluency1:tested_object1          -0.06149    0.20991  -0.293   0.7696
cognate1:fluency1:tested_object1 -0.69711    0.41958  -1.661   0.0966 .
```

3.3.2 Gaze duration data from the study phase.

The analysis of the eye-tracking data showed that the eyes of the participants did follow the audio across the screen. Figure 13 shows how the eyes moved around the screen in the scenes for Object 2 (i.e. the hesitation occurred between Object 1 and Object 2 in the dysfluent versions) in the different conditions. There is a clear fixation point on all four of the objects in the order of mention in the sound file. The grey graph which shows fixation on Object 2 spikes more than that of the other three graphs in the beginning of the scene observation, which could mean that, generally, Object 2 was the most visually dominating object in the scenes. The grey and the red graphs have a less dramatic high point in the two noncognate scenes than the black and pink graphs, and the fixation duration in the noncognate dysfluent scene is longer than in the noncognate fluent scene.

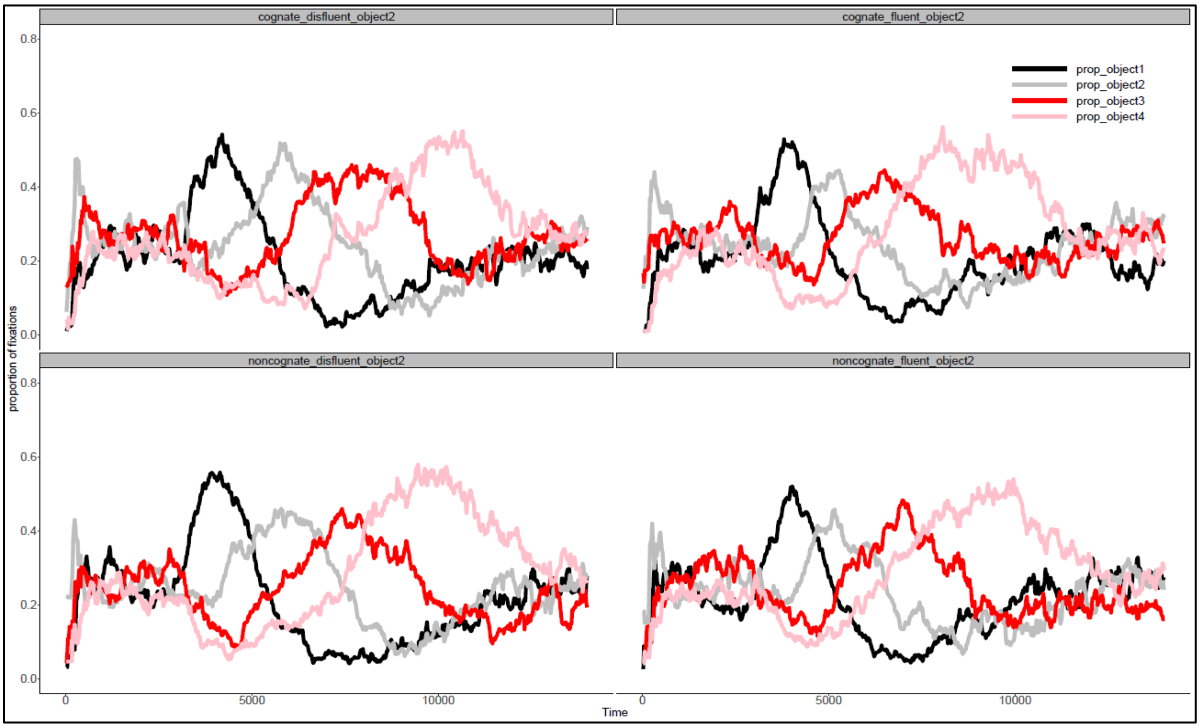

Figure 13. Eye tracking data in all four conditions.

The effects of dysfluency gaze durations (i.e. the total amount of time an object was looked at) is further depicted in Figure 14 which shows fixation times on Object 2 and 3 in the four conditions. The grey and pink graphs demonstrate longer fixation times when listening to dysfluent sentences than to fluent sentences. The best fitting linear mixed effects models of the gaze durations for Object 2 (Table 7) and Object 3 (Table 8) both show significant effects of fluency on dwell time.

Object 2 Object 3

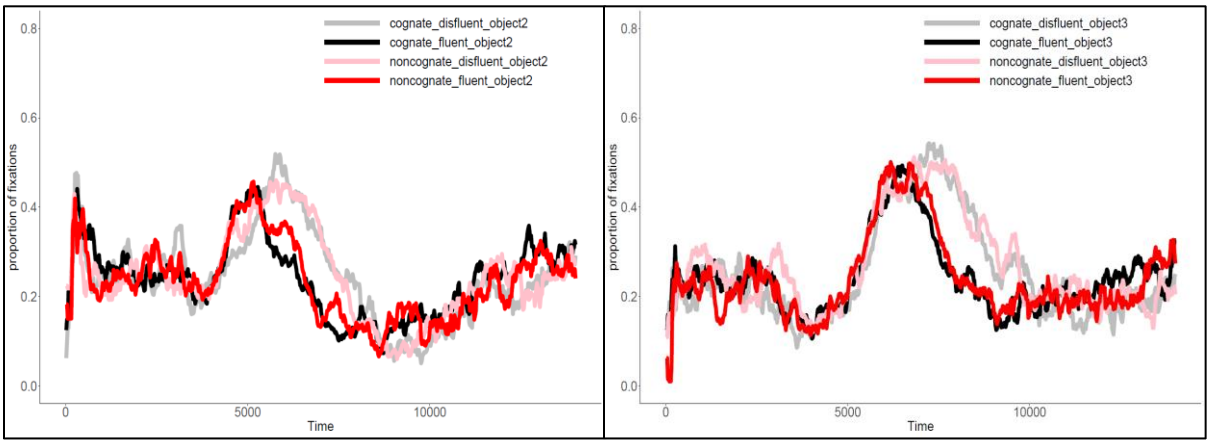

Figure 14. Eye tracking data with the four conditions in respectively object 2 and object 3.

Table 7. Best fitting model predicting dwell time on Object 2

	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	2055.47	127.22	42.97	16.157	< 2e-16	***
object_name1	-22.75	82.84	36.62	-0.275	0.78518	
fluency1	-170.83	55.79	1161.17	-3.062	0.00225	**
tested_object_freq_z	-27.04	54.07	43.32	-0.500	0.61957	
object_name1:fluency1	-82.04	111.36	1167.64	-0.737	0.46145	
object_name1:tested_object_freq_z	-91.79	114.81	45.88	-0.799	0.42814	
fluency1:tested_object_freq_z	68.91	55.04	1149.78	1.252	0.21088	
object_name1:fluency1:tested_object_freq_z	-126.41	109.96	1144.10	-1.150	0.25053	

Table 8. Best fitting model predicting dwell time on Object 3

	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	2146.5652	108.5479	49.6749	19.775	< 2e-16	***
object_name1	96.5206	74.1187	37.2134	1.302	0.201	
fluency1	-238.8608	57.6962	1134.9902	-4.140	3.73e-05	***
tested_object_freq_z	0.3276	51.5707	48.1995	0.006	0.995	
object_name1:fluency1	-69.9749	115.2064	1139.2598	-0.607	0.544	
object_name1:tested_object_freq_z	-103.1483	95.6160	44.2689	-1.079	0.287	
fluency1:tested_object_freq_z	86.0204	58.1309	1140.8704	1.480	0.139	
object_name1:fluency1:tested_object_freq_z	67.5533	116.1369	1138.6852	0.582	0.561	

3.3.3 Effects of gaze duration on accuracy in the test phase.

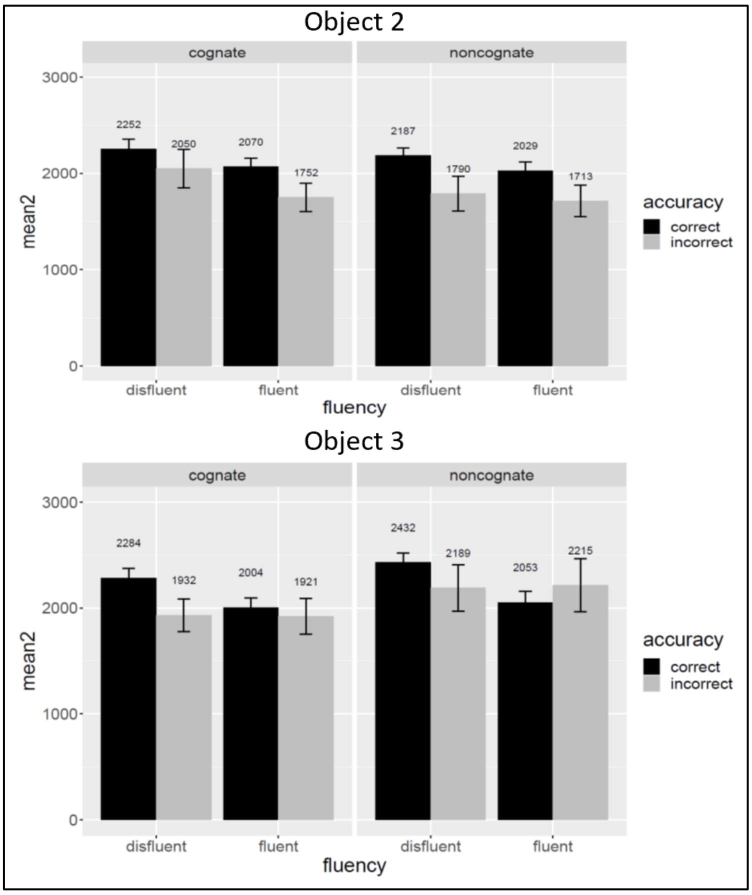

Figure 15. Mean dwell time in correct and incorrect answers in the study phase. Object 2 and object 3.

The following analysis tested whether dwell times on Objects 2 and 3 during the study phase influenced accuracy during the test phase. The mean dwell times for Objects 2 and 3 during the study phase are shown in Figure 14 for correct and incorrect answers in the test phase. The black columns show the mean dwell time of the correct answers in all four conditions and the grey columns show the mean dwell times of incorrect answers in all four conditions. Object 2 had a mean gaze duration of 2131.75 in correct responses and 1826.25 in incorrect responses. The correct answers have longer fixation times in Object 2 under all conditions. The pattern for Object 3 is less consistent, with the fluent noncognate condition showing shorter dwell times on the correct answers than on the incorrect answers. The dysfluent accurate responses in all but the noncognate fluent condition in Object 3 show longer gaze dwell times in the correct noncognate response than in the correct cognate responses.

The best fitting LME models are shown for Object 2 in Table 9 and Object 3 in Table 10 on the next page. The interaction of Fluency and Dwell time on Object 3 approached significance.

Table 9, Best fitting model predicting dwell time by accuracy for object 2

```
 Estimate Std. Error z value Pr(>|z|)
(Intercept)                        -1.62389    0.16997  -9.554   <2e-16 ***
object_name1                        0.05353    0.19816   0.270    0.787
fluency1                            0.19873    0.15344   1.295    0.195
dwell2_centered                    -0.16322    0.11371  -1.435    0.151
object_name1:fluency1               0.45490    0.30649   1.484    0.138
object_name1:dwell2_centered        0.13125    0.20121   0.652    0.514
fluency1:dwell2_centered           -0.05672    0.16114  -0.352    0.725
object_name1:fluency1:dwell2_centered -0.11787    0.32052  -0.368    0.713
```

Table 10, Best fitting model predicting dwell time by accuracy for object 3

```
 Estimate Std. Error z value Pr(>|z|)
(Intercept)                        -1.57362    0.14499 -10.853   <2e-16
***
object_name1                       -0.18066    0.19801  -0.912   0.3616
fluency1                            0.06487    0.15291   0.424   0.6714
dwell3_centered                    -0.10934    0.12901  -0.848   0.3967
object_name1:fluency1              -0.27392    0.30794  -0.890   0.3737
object_name1:dwell3_centered        0.12954    0.18994   0.682   0.4952
fluency1:dwell3_centered            0.30258    0.16135   1.875   0.0608
object_name1:fluency1:dwell3_centered -0.14432    0.32400  -0.445   0.6560
```

4 General discussion

The aim of our study was to see if cognate status and dysfluencies in L2 verbal input affected bilingual visual memory and processing. Our study is based on existing research on the effects of cognates in bilinguals and hesitation effects found in monolinguals but not in bilinguals (A. Konopka, in print). We wanted to see if highly proficient Norwegian-English bilinguals with two languages that share a vast number of cognates would be affected by cognate/noncognate status and if they, with high proficiency levels, would behave similarly to monolinguals (Corley et al., 2007) when exposed to dysfluencies, or if dysfluencies might not have any effect in L2 processing. In order to check proficiency levels, four shorter language proficiency tests were conducted, designed to collect objective data on proficiency levels, and a LEAP-Q was filled out to map self-rated subjective proficiency. The mapping was necessary when looking for a link between proficiency levels and performance in the visual memory test. The results are based on accuracy in the test phase and eye tracking data from the study phase, and correlations between accuracy, cognate status, fluency status and gaze durations. Our aim was first to determine if accuracy would increase when the object changes were on cognate objects versus noncognate objects and secondly, if dysfluent conditions would make recollection for either cognate, noncognate or both conditions better than fluent conditions. Lastly, because the experiment was eye tracked, we were curious to see if there were any eye movement effects made by cognate status and hesitations, and if these correlations also linked back to accuracy scores in the test phase.

In this discussion, I will relate our results to the findings of Corley et al. (2007), Sampaio & Konopka (2013) and the inside report from the in-print study by A. Konopka while discussing the results in relation to the four hypotheses mentioned in the present study section. I will also discuss how our findings and lack of findings give implications for models of bilingual processing and how this study proposes new questions for future research.

4.1 The LEAP-Q and language proficiency tests

The LEAP-Q responses show that our group of participants have high proficiency levels, and that the test group is uniform to a large degree even though there were some range on their proficiency ratings. Some of the readings were unusual. Results from the self-ranking show that some participants ranked their Norwegian proficiency as lower than one would expect from native speakers (e.g. range from 5-10 in speaking proficiency). This is probably culturally motivated, as overconfidence is considered a negative personal trait in Norway. This could have affected the Norwegian participants in their self-ratings and made them distrust and underestimate their own abilities. However, despite the possible underestimations in the participants' Norwegian proficiency, the correlations between the English subjective proficiency ratings, the objective English vocabulary test and particularly the LexTALE test were strongly suggestive that the participants rated their English skills more accurately than their Norwegian skills.

Unfortunately, none of the proficiency data interacted with any of our experimental findings. No correlations were found between the achievements in the language tests, the self-ranked proficiency and the visual memory test phase results. The individual scores and ratings were similar (even though the test results in the Norwegian and English vocabulary tests were somewhat varied). The LEAP-Q and the four proficiency tests confirmed that the test group were highly proficient Norwegian-English bilinguals.

4.2 The visual memory test

The main experiment in this study is the visual memory test where we manipulated cognate status and hesitations to see if this had any effects on visual scene memory. Cognate status and hesitations are the main manipulators, and in this section, I will go through the findings related to cognates and hesitations and their effects on memory, gaze durations and each other respectively. The results provide suggestions for the hypotheses mentioned in the present study (section) and give further implications for the reviewed research and models of bilingual processing in the introduction of this paper. The main factor we looked at was how cognate status and dysfluencies affected visual memory. The findings were not conclusive, but our data provides trends that could become significant in a larger test group. The rest of the discussion is based on the trends suggested by the results and not on compelling evidence.

4.2.1 Cognate and hesitation effects on memory

Results on accuracy showed little to no effect for cognates on memory. Object 2 (i.e. the object mentioned second in the scene) had a marginally higher mean accuracy in cognate conditions with more accurate responses in fluent conditions whereas Object 3 showed a marginally higher accuracy in the noncognate condition. In summary, no significant cognate effect on memory was found. There was also no single hesitation effect on memory independent of cognate status, but the three-way interaction between Cognate status, Fluency and Accuracy will be further discussed in section 4.2.3 below.

Hypothesis 1 predicted a better memory for cognates than noncognates, but no indication of this was found in the results. Furthermore, hesitations did not provoke better memory on its own independent of cognate status, and we did not find evidence for better memory in cognate conditions. Cognate status did not affect the eye movements of the participants, and the eye tracker confirmed that eye movements were consistent with the audio independent of cognate status and hesitations. So far, results indicate that the null hypothesis is more correct in its assumption of no cognate effects on memory, and that hesitations alone do not cause better memory.

Nonselective models of bilingual language processing propose that both languages are activated during bilingual processing, and evidence from cognates and interlingual homographs indicate that cognate status affects processing. Additionally, Sampaio & Konopka (2013) found that bilinguals remember surface form of words better than monolinguals. With these two findings combined we were curious to see if cognates, as a result of double activation in the mental lexicon, would be remembered better than noncognates with single activation, and whether, because of bilinguals' strong ability to remember surface form, this would show in accuracy measures or not. A possible explanation to the lack of findings supporting better memory for cognates could be that the bilinguals are helped by their ability to remember surface form in both cognate and noncognate conditions. However, this is difficult to test for as monolinguals with no cognate information in their lexicon cannot work as a control group.

Our findings are also in line with the results of Libben & Titone (2009) and Lagrou et al. (2013). The former investigated context effects in orthographic sentence comprehension and results proposed that highly proficient bilinguals experienced less cognate facilitation than the lower proficient participants in High-constraint contexts. Lagrou et al. (2013) tested sentence context in spoken word comprehension, and found evidence suggesting that high-constraint sentences were processed faster than low-constraint sentences, and that spoken words are recognised faster if the speaker is a native speaker of the listener's L2. One can argue that our sentences, as they were presented simultaneously with a limited amount of visual objects in a scene, were indeed highly constrained, and based on the findings of the two studies by Libben & Titone (2009) and Lagrou et al. (2013), the high-constraint sentences combined with the high levels of proficiency should lead to smaller cognate facilitation effects than if the participants were at lower proficiency levels.

Research on executive control suggests that bilinguals have better inhibitory- and attention control than monolinguals (Bialystok, 2011; Costa et al., 2009). Furthermore, evidence from the language mode hypothesis (Grosjean, 1988; Dunn & Fox Tree, 2014) could imply that highly proficient bilinguals who have a more balanced relationship between their two languages, as suggested by the RHM-model, enter a language mode where the two languages become more (but not exclusively) selective in activation. The Visual Memory test was only conducted in English, and with the sentences having the same structure, the highly proficient Norwegian-English bilinguals might have managed to control their attention towards one language (English) more so than both of their languages (Norwegian and English) while processing. However, cognates were predicted to disturb a monolingual language mode as they were hypothesised to force activation from the nontarget language due to the close links between cognates and shared concepts.

4.2.2 Cognate and hesitation effects on gaze durations

Results showed a clear gaze pattern following the audio in both cognate and noncognate conditions. There was no difference in the cognate scenes and noncognate scenes in terms of gaze durations. The graph in Figure 13 clearly shows that the participants' eye movements followed the pattern directed by the audio they listened to as they studied the scenes, with no differences in cognate and noncognate scenes in terms of eye movements. There was a slight tendency towards the suggestion that correct answers in the test phase had longer mean gaze durations in the study phase than incorrect answers, which proposes that there is a link between gaze duration and memory. Further, the interaction between Fluency and Gaze duration on accuracy in Object 3 approached significance in the analysis. This is not surprising as hesitations was hypothesised to maybe cause longer gaze durations whereas longer gaze durations were expected to maybe facilitate memory. However, I will not focus on the gaze duration effects on memory, although the plausible link between fluency and gaze durations on memory later in the sentence is an interesting observation. As the language tests suggest, the participants are all very proficient in English, and the results indicate that both cognate (fluent/dysfluent) and noncognate (fluent/dysfluent) conditions trigger eye movements to follow the verbal input in highly proficient bilinguals. In a larger group of participants, definite differences in the cognate/noncognate conditions in terms of eye movement patterns and gaze durations might have been seen.

Hypothesis 2 predicted longer gaze durations for cognates than noncognates, but no gaze duration differences were found between the two conditions. The second prediction; that hesitations should provoke longer gaze durations on the following object in high-proficient bilinguals was, on the other hand, confirmed. Our eye tracking data strongly suggests that hesitations affect gaze durations on the object following, independent of cognate/noncognate status, and Object 3 also showed an effect of dysfluent conditions on accuracy in both cognate and noncognate versions.

Assumptions by the language mode hypothesis relating to variance in language selectivity combined with the IC-model's assumptions of inhibition- and attention control could suggest that the English-only testing could have provoked a more selective language processing in this study. If a language mode was achieved, the bilinguals might have been able to focus more exclusively on the ENGLISH tags (as proposed by the IC model) and been better at inhibiting the NORWEGIAN tagged words. This would cause a lesser amount of words in competition, given the assumption that a language mode might make them able to not take NORWEGIAN into account as much as they would if they processed in a language switching context. The IC model is not a pure language nonselective model as it implies that the Supervisory Attentional System (SAS) to a large extent controls what activates and not through the task schemas. Our eye tracking results could imply that the SAS and the language task schemas become more closely linked in bilinguals when they process exclusively in one of their two languages (if they are highly proficient).

The RHM model (Kroll & Stewart, 1994) is mainly a production model that proposes that links between the L1 and L2 is weaker than the links from L2 to L1, and that L1 has stronger links to concepts than does L2 (see Figure 4 in section 1.3). This model focuses on proficiency, and how links between L1, L2 and concepts strengthen with rising levels of proficiency. A balanced bilingual will thus have balanced L1 and L2 with equal links to each other and concepts. Obviously, the bilinguals we tested had strong links between their L1 and L2 and between L2 and concepts, but they were not balanced bilinguals. Because L1 was more dominant than L2 in all the participants, it took more focus to inhibit L1 and focus on L2 that it would have the other way around. When L1 is heavily suppressed, it does not interfere much in processing, and it takes people longer to retrieve lexical items from L1 after L2 processing than it does to retrieve L2 after L1 processing. This is probably due to the extreme suppression it takes to subdue the more dominant language compared to the least dominant language, and the effort it takes to retrieve the L1 from the suppressed state. Again, as with the IC model and the BIA+ model, the eye tracking results might suggest that the link between L1 and L2 is absent in highly proficient bilinguals who only process one of their languages over time. Our results could suggest that in L2-exclusive production in highly proficient bilinguals, the links between L2 to L1 are weaker than the links from L1 to L2 as a result of the suppression of L1 in processing.

4.2.3 Cognate effects on dysfluencies

Results show an effect on gaze durations after hesitations, but no difference was found between cognate and noncognate conditions. In total, the dysfluent gaze durations on Object 2 were 4302 for cognates and 3977 for noncognates, whereas the fluent gaze durations were 3822 for cognates and 3742 for noncognates. Object 3 also suggested the same, that dysfluencies cause some longer gaze durations to the preceding object both in cognate and noncognate conditions. An observation which is closer to significance than the previously mentioned data is the link between Dysfluency, Noncognate and Object 2 (but not Object 3). Figure 12 implies higher accuracy when the dysfluent noncognate condition comes early in the scene, whereas the results are inconclusive in the same condition later in the sentence (i.e. Object 3).

Hypothesis 3 predicted the possibility of high-proficient bilinguals to experience hesitation effects on cognates and noncognates, but it also assumed that there would not be a difference in effects between the two different cognate conditions. However, we were surprised to find trends to suggest that noncognates are better remembered when followed by a hesitation in early processing. The internal report from A. Konopka stated that they did not find evidence for hesitation effects in bilinguals as Corley et al. (2007) found in monolinguals, and our results propose that hesitation effects might depend on proficiency levels. It is possible that, because cognates cause double activation, there is no more room for extra activation as they already activate more than noncognates, which on the other hand, can afford extra activation from dysfluencies. When processing noncognates, two different words linked to the same concept compete for selection (according to nonselective models), and it is possible that the extra attention provoked by hesitations strengthens focus through lexical selection with noncognates. One would, however, expect to see hesitation effects in both conditions, and with a larger test group it would be interesting to see if cognates also showed greater accuracy in dysfluent conditions.

Corley et al. (2007) found effects of hesitations both short-term and long-term, and their results were replicated to some degree. The short-term effect showed longer gaze durations on the object mentioned after a hesitation, and noncognate+dysfluency showed a long-term memory effect in early sentence processing. Corley et al. (2007) propose that a hesitation makes the language user pay more attention to the next word mentioned, and our eye tracking data shows that hesitations also make bilinguals more attentive of the following object. However, we did not find a clear long-term memory effect from the hesitations in cognate objects.

Why we found different trends in Object 2 versus Object 3 remains the elephant in the room. We did not expect to see different implications in the two objects when we designed the experiment. It seems that early sentence processing in L2 behaved similarly to monolinguals (Corley et al., 2007) unlike late sentence processing which did not imply dysfluency facilitation effects on accuracy. Processing in one's second language is difficult, and much effort is required to suppress and inhibit L1 to retrieve L2. It is possible that highly proficient bilinguals get more delayed gradually in processing throughout a sentence, where the delay is small in the beginning of the sentence but then adds up as more words need decoding. Our results imply that there might be a retrieval lag that becomes more severe the further one gets into longer utterances. To summarise, we found short-term effects of dysfluencies on gaze durations in both cognate and noncognate conditions, and long-term effects of dysfluencies in early noncognate conditions. No cognate effect on visual memory was found in this study, but the question as to why we did not achieve a cognate effect still remains.

4.2.4 Limitations and future research

This study was conducted on 33 participants, and the main limitation is that the study is underpowered. As the stimuli were divided into eight lists, only four participants were tested per list. If the mirrored versions are not considered, this means that only eight people were tested in each condition which is not enough for solid and reliable results. The findings discussed above are thus only borderline effects from the data, and no conclusion can be drawn on cognate effects on visual memory or the lack of link between cognate and fluency status until more participants are tested in each condition. In addition to the lack of power, the study only tested bilinguals with high proficiency levels. This is a strength in that our results were found from a uniform test group, but it also limits the findings as we do not know if the results are caused by bilingualism as predicted by Sampaio & Konopka (2013), high proficiency levels or a combination of both. As only bilinguals were tested, we do not yet know if the lack of dysfluency effect in noncognate conditions late in sentence processing is a result of retrieval lag in bilinguals or a general effect that can be seen in both monolinguals and bilinguals due to general limitations in memory. To investigate the possible retrieval lag-effect, one would need for the effect to disappear when testing monolinguals as they should not have a delay in sentence processing.

One of the main questions relating to our results is why we did not find any dysfluency effects in cognate conditions on Object 2. Pictures of cognates are used mostly in production studies, and when using cognate pictures in comprehension, we were limited to cognates which could be placed logically in a scene. Because the objects had to be concrete nouns which are "simple" words, we could not include more challenging abstract nouns in the stimuli. It is possible that, as a result of the concrete noun restriction, we did not find cognate effects as the words were easy to process with a high general frequency. Unfortunately, it would not be possible in this type of experiment to include abstract objects as they could not be illustrated in the visual scenes.

This study is the first of its kind. No other study has looked at cognate and dysfluency effects on memory for visual scenes in L2 processing of highly proficient bilinguals, and our results show some interesting trends in the data that provide insights for what is necessary to investigate in future research on the topic. More research is needed on several aspects to determine the reasons for the observations made in this study:

- A larger test group of Norwegian-English bilinguals in the same conditions is necessary to more accurately determine the effects of cognates on memory and the link between cognate status and hesitations on the effect on visual scene memory.
- The same experiment should be conducted on low-proficient Norwegian-English bilinguals to see whether cognate facilitation on visual scene memory and cognate dysfluency effects rely on proficiency levels.
- The experiment should be conducted in mixed trials to test for language mode effects and language selectivity in highly proficient bilinguals.
- The experiment should be conducted on monolinguals to determine if the difference in dysfluency effects in noncognates on Object 2 and Object 3 is a result of bilingual retrieval lag, or if it is only a result of too much input.

5 Conclusion

The current study investigates the possible effects of cognates and dysfluencies on memory for visual scenes in the L2 processing of highly proficient Norwegian-English bilinguals. In conclusion, we did not find any cognate effects on visual scene memory as was expected based on nonselective processing theory and previous findings by Sampaio & Konopka (2013). This study is the first of its kind to allow for cognate manipulation to test for memory effects. Sampaio & Konopka (2013) investigated the difference in monolinguals' and bilinguals' retrieval of surface form, where bilinguals remembered surface form better than monolinguals, and our findings are arguably in line with the results of that study as the surface forms were remembered well despite cognate status. Hesitations have proven to affect attention on what follows (Corley et al. 2007), but this effect was not confirmed in bilinguals in a recent study by Konopka (internal report from an in-print study), however, we found hesitation effects on gaze durations in both cognate and noncognate conditions as well as higher accuracy on noncognate objects following a hesitation early in sentence processing (Object 2). Our participants behaved more similarly to monolinguals than expected, which propose that hesitation effects in bilinguals depend on level of proficiency. The interesting indication in our results is the suggestion that retrieval lag might prevent the bilinguals from monolingual-like processing of hesitations appearing later in sentences, but this needs further exploration.

Overall, more research is necessary to determine the possible effects of cognates and hesitations on bilingual visual memory both on different proficiency levels and monolinguals. Our results imply that more highly proficient bilinguals might process visual and verbal input similarly to monolinguals at the beginning of sentences, but that a retrieval lag might cause bilinguals to fall behind as the input exceeds a certain length.

6 References

Adesope, O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research, 80*, 207-245.

Baumgart, C. Q., & Billick, S. B. (2018). Positive cognitive effects of bilingualism and multilingualism on cerebral function: a review. *Psychiatric Quarterly, 89*(2), 273-283. DOI: 10.1007/s11126-017-9532-9

Bialystok, E. (2009). Bilingualism: The good, the bad and the indifferent. *Bilingualism: Language and Cognition, 12*(1), 3-11. DOI: 10.1017/S1366728908003477

Bialystok, E. (2011). Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology, 65*(4), 229-235. DOI: 10.1037/a0025406

Bialystok, E., Craik, F. I. M., & Freedman, M. (2007). Bilingualism as protection against the onset of symptoms of dementia. *Neuropsychologia, 45*, 459-464.

Bialystok, E., Craik, F. I. M., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition, 34*, 858-873. DOI: 10.1037/0278-7393.34.4.859

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science*, *7*(3), 325-339.

Boersma, P., & Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.1.05. Retrieved 01 November 2019 from http://www.praat.org/

Campbell, R., & Sais, E. (1995). Accelerated metalinguistic (phonological) awareness in bilingual children. *British Journal of Developmental Psychology, 13*, 61–68.

Cat, n.1. (2020). In OED Online. Oxford University Press. Accessed April 15, 2020, from https://oed.com/view/Entry/28649?rskey=aYEoMO&result=2&isAdvanced=false#eid

Corley, M., MacGregor, L. J., & Donaldson, D. I. (2007). It's the way you, er, say it: Hesitations in speech affect language comprehension. *Cognition*, *105*(3), 658-668. DOI: https://doi.org/10.1016/j.cognition.2006.10.010

Costa, A., Caramazza, A., & Sebastián-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*, 1283-1296.

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition 113*, 135-149.

Costa, A., Pannunzi, M., Deco, G., & Pickering, M. J. (2017). Do bilinguals automatically activate their native language when they are not using it? *Cognitive Science*, *41*, 1629-1644.

De Groot, A. M., & Nas, G. L. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30*(1), 90–123. https://doi.org/10.1016/0749-596X(91)90012-9

Diamond A. (2013). Executive functions. *Annual review of psychology*, *64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*(4), 496–518. https://doi.org/10.1006/jmla.1999.2654

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism; Cambridge 5*(3), 175-197

Dijkstra, T., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: effects of task demands and language intermixing. *Bilingualism: Language and Cognition*, *1*(1), 51–66. https://doi.org/10.1017/S1366728998000121

Dunn, A. L., & Fox Tree, J. E. (2014). More on language mode. *International Journal of Bilingualism, 18*(6), 605-613.

Ecke, P. (2004). Language attrition and theories of forgetting: A cross-disciplinary review. *International Journal of Bilingualism*, *8*(3), 321-354.

EF English Proficiency Index. (2019). *The World's Largest Ranking of Countries by English Skills.* Retrieved from https://www.ef.no/epi/regions/europe/norway/

Flavell, J. H. (1978). Metacognitive development. In J. M. Scandura & C. J. Brainerd (Eds.), *Structural process theories of complex human behaviour*, 213–245.

Gollan, T., & Acenas, LA. (2004). What Is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30,* 246-69. 10.1037/0278-7393.30.1.246.

Gollan, T. H., Salmon, D. P., Montoya, R. I., & Galasko, D. R. (2011). Degree of bilingualism predicts age of diagnosis of Alzheimer's disease in low-education but not in highly educated Hispanics. *Neuropsychologia*, *49*(14), 3826–3830.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67–81. https://doi.org/10.1017/S1366728998000133

Green, D., & Abutalebi, J. (2013) Language control in bilinguals: the adaptive control hypothesis. *Journal of Cognitive Psychology, 25*, 515-530.

Grosjean, E. (1998). Studying bilinguals; methodological and conceptual issues. *Bilingual Language and Cognition, 1*, 131-149.

Harbert, W. (2007). *The Germanic Languages.* New York: Cambridge University Press.

Hernandez, A. E., & Li, P. (2007). Age of acquisition: its neural and computational mechanisms. *Psychological bulletin*, *133*(4), 638–650. https://doi.org/10.1037/0033-2909.133.4.638

Kaushanskaya, M. & Marian, V. (2007). Bilingual language processing and interference in bilinguals: Evidence from eye tracking and picture naming. *Language Learning, 57,* 119–163.

Kemp, C. (2007). Strategic processing in grammar learning: Do multilinguals use more strategies? *International Journal of Bilingualism*, *4*, 241–261.

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, *9*(2), 119–135. https://doi.org/10.1017/S1366728906002483

Kroll, J. F., & De Groot, A. M. B. (Edd.). (2005). *Handbook of Bilingualism: Psycholinguistic approaches.* New York: Oxford University Press.

Kroll, J. F., Michael, E., Tokowicz, N., & Dufour, R. (2002). The development of lexical fluency in a second language. *Second Language Research*, *18*, 137-171.

Kroll J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetrical connections between bilingual memory representations. *Journal of Memory and Language*, *33*(2), 149-174. https://doi.org/10.1006/jmla.1994.1008

Kroll, J. F., van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The revised hierarchical model: A critical review and assessment. *Bilingualism: Language and Cognition, 13*(3), 373-381.

Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2013). The influence of sentence context and accented speech on lexical access in second-language auditory word recognition. *Bilingualism: Language and Cognition, 16*(3), 508-517.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behaior Research Methods, 44,* 325-343.

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(2), 381–390. https://doi.org/10.1037/a0014875

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and monolinguals. *Journal of Speech, Language and Hearing Research*, *50*(4), 940-967. https://doi.org/10.1044/1092-4388(2007/067)

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324.

Matthews, P. (2014). bilingual. In The Concise Oxford Dictionary of Linguistics (3 ed.). Oxford: Oxford University Press. Retrieved 09. April 2020, from https://www.oxfordreference.com/view/10.1093/acref/9780199675128.001.0001/acref-9780199675128-e-365?rskey=FgGII0&result=7.

Merriam-Webster. (n.d.). Executive control. In *Merriam-Webster.com dictionary*. Retrieved May 10, 2020, from https://www.merriam-webster.com/dictionary/executive%20control

Michael, E. B., & Gollan, T. H. (2005). Being and becoming bilingual: Individual differences and consequences for language production. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (p. 389–407). Oxford University Press.

Moon, n. (2020). In OED Online. Oxford University Press. Accessed April 22, 2020, from https://oed.com/view/Entry/121900?rskey=muzXZA&result=1&isAdvanced=false#eid

New American Economy. (2017). *Not lost in translation: The growing importance of foreign language skills in the U.S. job market.* Retrieved April 10, 2020 from http://www.newamericaneconomy.org/wp-content/uploads/2017/03/NAE_Bilingual_V9.pdf

One, adj., n., and pron. (2020). In OED Online. Oxford University Press. Accessed April 22, 2020, from https://oed.com/view/Entry/131358?rskey=LNudsd&result=1&isAdvanced=false#eid

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., & Fazio, F. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: An fMRI study during verbal fluency. *Human Brain Mapping, 19*, 170–182.

Portocarrero, J. S., Burright, R. G. & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology, 22*, 415–422.

Reindeer, n. (2020). In OED Online. Oxford University Press. Accessed April 22, 2020, from https://oed.com/view/Entry/161586?redirectedFrom=reindeer#eid

Sampaio, C., & Konopka, A. E. (2013). Memory for non-native language: The role of lexical processing in the retention of surface form. *Memory, 21*(4), 537-544. https://doi.org/10.1080/09658211.2012.746371

Sánchez-Casas, R. & García-Albea, J. E. (2005). The representation of cognate and noncognate words in bilingual memory. In J. Kroll & A. DeGroot (Ed.), *Handbook of Bilingualism* (226-250). New York: Oxford University Press.

Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition, 13*(2), 231-252.

Schröter, P. & Schroeder, S. (2016). Orthographic processing in balanced bilingual children: cross-language evidence from cognates and false friends. *Journal of Experimental Child Psychology, 141*, 239-246. https://doi.org/10.1016/j.jecp.2015.09.005

Schweiter, J., & Sunderman, G. (2009). Concept selection and developmental effects in bilingual speech production. *Language Learning, 59*, 897-927.

Star, n. (2020). In OED Online. Oxford University Press. Accessed April 22, 2020, from https://oed.com/view/Entry/189081?rskey=csMGtk&result=1&isAdvanced=false#eid

Stern, Y., Habeck, C., Moeller, J., Scarmeas, N., Anderson, K. N., Hilton, H. J., Flynn, J., Sackeim, H., & van Heertum, R. (2005). Brain networks associated with cognitive reserve inn healthy young and old adults. *Cerebral Cortex, 15*, 394– 402.

Sun, n. (2020). In OED Online. Oxford University Press. Accessed May 29, 2020, from https://www.oed.com/search?searchType=dictionary&q=sun&_searchBtn=Search

Sunderman, G., & Kroll, J. (2006). First language activation during second language lexical processing: An investigation of lexical form, meaning, and grammatical class. *Studies in Second Language Acquisition*, *28,* 387 - 422. 10.1017/S0272263106060177.

Take, v. (2020). In OED Online. Oxford University Press. Accessed April 22, 2020, from https://oed.com/view/Entry/197158?rskey=FUATZW&result=3&isAdvanced=false#eid

That, pron.1, adj., and adv. (2020). In OED Online. Oxford University Press. Accessed April 22, 2020, from https://oed.com/view/Entry/200177?rskey=hkxbfw&result=1&isAdvanced=false#eid

They, pron., adj., adv., and n. (2020). In OED Online. Oxford University Press. Accessed April 22, 2020, from https://oed.com/view/Entry/200700?redirectedFrom=they#eid

Three, adj., and n. (2020). In OED Online. Accessed April 22, 2020, from https://oed.com/view/Entry/201165?rskey=SE3C4f&result=1&isAdvanced=false#eid

Two, adj., n., and adv. (2020). In OED Online. Oxford University Press. Accessed April 22, 2020, from https://oed.com/view/Entry/201165?rskey=SE3C4f&result=1&isAdvanced=false#eid

Van Hell, J.G., & Tanner, D. (2012), Second language proficiency and cross-language lexical activation. *Language Learning*, *62,* 148-171.

Van Heuven, W.J.B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology, 67*, 1176-1190.

Vorstman, E., De Swart, H., Ceginskas, V., & Van Den Bergh, H. (2009). Language learning experience in school context and metacognitive awareness of multilingual children. *International Journal of Multilingualism*, *6,* 258–280.

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*. Translated by Ogden, C. K. London: Routledge & Kegan Paul.

Yu, Z., & Schwieter, J. W. (2018). Recognising the effects of language mode on the cognitive advantages of bilingualism. *Frontiers in Psychology, 9*, 1-6.

7 Appendices

Appendix A- Consent form

PARTICIPANT INFORMATION SHEET AND CONSENT FORM

INVITATION TO PARTICIPATE IN A RESEARCH PROJECT

MEMORY PROCESSES IN BILINGUALS

We are looking for native speakers of Norwegian to take part in a language study investigating memory processes in bilinguals.

In order to participate in this study you need to be a **native speaker** of Norwegian with no other home languages (excluding perhaps English) and have a reasonable proficiency in English as your second language. You should have normal or corrected-to-normal vision (including colour vision) and hearing and have no diagnosed language impairments such as dyslexia or stuttering.

The study has three main components:

1. A language background questionnaire
2. Some short language proficiency tasks
3. A picture memory task with eye-tracking

Completeing all tasks will take around 2 hours. The study is run by MA students Andrea Skintveit Holter (andrea@skintveit.com) and Irene Slethei (ireneslethei@hotmail.com). Please contact them if you have any queries about the study. The research is supervised by Professor Linda Wheeldon (linda.r.wheeldon@uia.no) and Professor Allison Wetterlin (allison.wetterlin@uia.no).

WHAT IS THE STUDY ABOUT?

This study is designed to investigate aspects memory when using English as a second language. The three components which will be completed English:

1. A questionnaire asking questions about your language background and about how you rate your own level of proficiency in different aspects of the languages that you speak. It should take about 40 minutes to complete.
2. Simple language tests: one assessing vocabulary and involving listening to nonsense words. These tests will take approximately 25 mins to complete.
3. A memory task involving descriptions of pictured scenes. You will first see the scenes and and hear English descriptions of them (30 mins). After some mathematics you will be asked to decide which pictures you have seen before. Your eye-movements will be recorded during this task using non-invasive eye-tracking technology.

If, after having read the information below, you decide to take part in the study please complete the consent form at the end of this document.

The study will collect and record personal information about you. However, all your data will be pooled with that of other participants for statistical analysis. You will never at any time be mentioned as an individual in relation to this study. Your personal data will be assigned a number code related to your name and stored on a non-networked, password protected PC. Only the laboratory directors and experimenters will have access to the key relating your data number to your name. In addition, we will record the responses you produce during the experiment, this includes key strokes and eye-movements. These data will be also be anonymised and treated as described above.

VOLUNTARY PARTICIPATION AND THE POSSIBLITY TO WITHDRAW CONSENT (OPT-OUT)

Participation in the study is voluntary. If you wish to take part, you will need to sign the declaration of consent on the last page of this document. This will allow us to process your data. You can, at any given time and without reason withdraw your consent. If you decide to withdraw participation in the project, you can ask that your test results and personal data be deleted, unless the data and tests have already been analysed or used in scientific publications.

So long as you can be identified in the collected data you have the right to:

- access the personal data that is being processed about you
- request that your personal data is deleted
- request that incorrect personal data about you is corrected/rectified
- receive a copy of your personal data (data portability), and
- send a complaint to the Data Protection Officer or The Norwegian Data Protection Authority regarding the processing of your personal data.

If you at a later point, wish to withdraw consent or have questions regarding the project, you can contact the project supervisors. Questions about the study or withdrawing consent can also be directed to the University og Agder's Data protection officer Ina Danielsen ina.danielsen@uia.no or NSD (Norsk senter for forskningsdata AS) by email personvernombudet@nsd.no or telephone 55 58 21 17.

WHAT WILL HAPPEN TO YOUR INFORMATION?

The information that is recorded about you will only be used as described in the purpose of the study.

All information is anonymised immediately and will be processed and used without your name or personal identification number, or any other information that is directly identifiable to you.

The results derived from the pooled data will be published. In the interest of being open to the scientific community and others interested in this research we would also like, with your permission, to publish the anonymised data to an open access database. If you agree to this, please sign under "publishing anonymised data to open access database" at the end of this document. The decision you make does not affect your eligibility for this study.

The principal investigator has the responsibility for the daily operations/running of this research project and that any information about you will be handled in a secure manner. Information about you is anonymised and your associated personal details will be deleted when the project is finished. The project period lasts until 20.12.2021.

FINANCE

In appreciation for your time and effort, you will receive a voucher for Sørbok/SiA for 150 NOK on completion of this study. No payment will be received for partial participation.

APPROVAL

Based on an agreement with The University of Agder, NSD – The Norwegian Centre for Research Data AS has assessed that the processing of personal data in this project is in accordance with data protection legislation.

CONSENT FOR PARTICIPATING IN THE RESEARCH PROJECT

I AM WILLING TO PARTICIPATE IN THE RESEARCH PROJECT

TITLE: MEMORY PROCESSES IN BILINGUALS

1) I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.
2) I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason. I understand that I can withdraw my data at any time during the experiment and after completion of the study until the data is analysed.
3) I understand that data collected during the study will be looked at by researchers from the University of Agder. I give permission for these individuals to have access to my data and to use it for research purposes. I understand that my data will be stored anonymously.
4) I agree to take part in the study.

Date Participant's Signature

Participant's Name (in BLOCK LETTERS)

PUBLISHING ANONYMISED DATA TO OPEN ACCESS DATABASE

I confirm that anonymised data can be uploaded to an open access database.

---

date Participant's Signature

Appendix B- Revised LEAP-Q Questions

Screening:

1. What is your age?
2. What is your gender?
3. Are you a native speaker of Norwegian?
4. Is Norwegian the only language you speak at home (aside from perhaps English)?
5. Are you a reasonably good speaker of English?
6. Do you have normal vision or vision that is corrected to normal with glasses or contact lenses?
7. Can you confirm that you have no language impairments such as dyslexia, stuttering etc.?
8. Do you have normal hearing or hearing that is corrected to normal?
9. Are you left or right handed?
10. What is country of birth?
11. What is your current country of residence?
12. How many years of education do you have?
13. What is the highest education level you have?
14. Have you participate in any experiments here before?

Language background:

1. Please list all the languages you speak in order of DOMINANCE (up to 5).
2. Please list all the languages you speak in order of ACQUISITION (up to 5).
3. Please list what percentage of the time you are on average exposed to each language (e.g. exposure in terms of talking, listening, and reading, including TV, films and music). (All your answers should add up to 100%)
4. Please list what percentage of the time you spend speaking each language. (All your answers should add up to 100%)
5. Please list what percent of time you typically spend reading in each language.
6. When choosing a language to speak, with a person who is equally fluent in all your languages, what percentage of time would you choose to speak each language? Please report percent of total time.
7. What cultures do you identify with (e.g., Norwegian, British, American, etc)? Please list each culture below (up to 5) and use the scale from 0-10 to rate the degree of identification, whereby 0 = no identification, 5 = moderate identification, 10 = complete identification.
8. Do you feel that you were once better in one of your languages and that you have become less fluent? If yes, which one? And at what age did you become less fluent?
9. In which language do you usually do the following tasks?
    - Simple maths (count, add)
    - Dream
    - Express anger or affection
    - Talk to yourself

Norwegian English Proficiency:

1. Please list the number of years and months you have spent in each language environment:

	Norwegian	English
A country where this language is spoken		
A family where this language is spoken		
A school where this language is spoken ALL of the time		
A school where this language is spoken SOME of the time		
A workplace where this language is spoken ALL of the time		

A workplace where this language is spoken SOME of the time		

2. Please rate how much the following factors contributed to your learning of each language on a scale of 0-10 whereby 0 = not a contributor, 5 = moderate contributor and 10 = most important contributor.

	Norwegian	English
Interacting with friends / colleagues		
Interacting with family		
Reading (e.g., books, magazines, online material)		
School and education		
Self-instruction (e.g., language learning videos or apps)		
Watching TV / streaming		
Listening to music/media		

3. Please rate to what extent you are currently (e.g. in the last month or so) exposed to each language on a scale of 0-10 whereby 0 = never, 5 = half of the time and 10 = almost always.

	Norwegian	English
Interacting with friends		
Interacting with family		
Reading (e.g., books, magazines, online material)		
Self-instruction (e.g., language learning videos or apps)		
Watching TV / streaming		
Listening to music/media		

4. Please rate your level of proficiency in the following aspects of each language on a scale of 0-10 whereby: 0 = none; 1 = very low; 2 = low; 3 = fair; 4 = slightly less than adequate; 5 = adequate; 6 = slightly more than adequate; 7 = good; 8 = very good; 9 = excellent; 10 = perfect.

	Norwegian	English
Speaking (general fluency)		
Pronunciation (accent)		
Reading		
Writing		
Grammar		
Vocabulary		
Spelling		

5. Please list the AGE (in years) you were when the following occurred for each language.

	Norwegian	English
Started hearing this language on a regular basis		
Became fluent in speaking this language		
Started learning to read in this language		
Became fluent in reading this language		

Dialect and Accent

1. Which dialect of Norwegian do you speak?

2. How important is speaking your own dialect for you on a scale of 0-10 (whereby 0 = not at all, 5 = moderately important, 10 = extremely important)?
3. To what extent would you say you modify your own dialect when speaking to a person with a different dialect on a scale of 0-10 (whereby 0 = not at all, 5 = moderately, 10 = totally)?
4. Have you lived in an environment where you have been exposed to other dialects than your own for a longer period of time (e.g. moving to a different city in Norway or living with someone who speaks another dialect)? If yes, which dialect? And for how long (in years)?
5. In your opinion how strongly regional is your spoken Norwegian on a scale of 0-10 (whereby 0 = not at all, 5 = moderately, 10 = very much)?
6. What kind of accent do you think your spoken English has (e.g., British / American / other / none in particular)?
7. In your view, how much of a Norwegian accent do you have when you speak English on a scale of 0-10? Whereby 0 = none, 1 = almost none, 2 = very light, 3 = light, 4 = some, 5 = moderate, 6 = considerable, 7 = heavy, 8 = very heavy, 9 = extremely heavy, 10 = pervasive.
8. To what extent do you think others identify you as a non-native speaker based on your ACCENT when speaking English on a scale of 0-10 (whereby 0 = never, 5 = half of the time 10 = always)?
9. How important is it for you to have a good accent when speaking English on a scale of 0-10 (whereby 0 = not at all, 5 = moderately important, 10 = extremely important)?
10. How much effort have you put into improving your accent when speaking English on a scale of 0-10 (whereby 0 = no effort at all, 5 = moderate effort, 10 = constant effort)?
11. How would you rate your ability to imitate foreign accents and dialects on a scale on a scale of 0-10 (whereby 0 = extremely poor, 5 = moderate, 10 = extremely good)?
12. Please rate the degree to which you agree with the following statements on a scale of 0-10 (whereby 0 = very strongly disagree, 10 = very strongly agree)?

Statement	Rating
It is important to me to speak grammatically correct English	
I pay attention to how people pronounce words and sounds	
I want to improve my pronunciation of English	
If it were possible I would like to pronounce English like a native speaker	
Pronunciation is not important to me because it does not affect how well I can communicate	

13. Are there any sounds in the English language you find difficult to pronounce?
   - If yes, which one(s)? (Write down the letter representing the sound or a word that contains the sound (capitalize the sound).

14. Have you noticed any English speech sounds that are difficult for other Norwegians when speaking English?
   - If yes, which one(s)? (Write down the letter representing the sound or a word that contains the sound (capitalize the sound).

15. When you are speaking do you ever find yourself accidentally mixing words or sentences from Norwegian and English?
   - If yes, how often does English accidentally intrude in your Norwegian on a scale of 0-10 (whereby 0 = never, 5 = half of the time, 10 = all of the time)

- And how often does Norwegian accidentally intrude into your English on a scale of 0-10 (whereby 0 = never, 5 = half of the time, 10 = all of the time)?

16. When you are speaking with a person who also knows both Norwegian and English do you ever find yourself intentionally mixing words or sentences from Norwegian and English
    - If yes, how often do you intentionally use English words when speaking Norwegian on a scale of 0-10 (whereby 0 = never, 5 = half of the time, 10 = all of the time)?
    - And how often do you intentionally use Norwegian words when speaking English on a scale of 0-10 (whereby 0 = never, 5 = half of the time, 10 = all of the time)?
17. Which written form of Norwegian have you predominantly been using?

Appendix C- LEAP-Q Selected Results

Question	Norwegian		English	
	Mean value	**Range**	**Mean value**	**Range**
Age	24.6	18-32		
Gender	19 F. 14 M			
Years of education	16.3	12.5-19		
Exposure (% of the time)	61.5	30-80	36.7	19-70
Speaking (% of the time)	84.2	40-99	14.3	1-40
Reading (% of the time)	49.7	10-95	50.3	5-90
Language of choice (% of the time)	83.8	0-100	14.6	0-100
Maths language (out of 33 participants)	32		1	
Dream language (out of 33 participants)	31		2	
Anger language (out of 33 participants)	31		2	
Self-talking language (out of 33 participants)	30		3	
Contributor- family (from 0-10)	9.4	3-10	2.5	0-10
Contributor- friends (from 0-10)	6.8	2-10	5.8	0-10
Contributor- reading (from 0-10)	6.9	3-10	7.8	4-10
Contributor- school (from 0-10)	7.9	3-10	7.5	2-10
Contributor- TV (visual) (from 0-10)	4.1	0-9	7.5	1-10
Contributor- music (audio) (from 0-10)	3.5	0-8	6.6	0-10
Exposure- family (from 0-10 of the time)	9.2	0-10	0.8	0-10
Exposure- friends (from 0-10 of the time)	8.3	5-10	2	0-6
Exposure- reading (from 0-10 of the time)	4.7	1-9	5.5	1-10
Exposure- TV (visual) (from 0-10 of the time)	3.2	0-7	7.1	3-10
Exposure- music (audio) (from 0-10 ott)	3	0-7	7.2	3-10
Vocabulary proficiency (self-rated from 1-10)	8.2	6-10	7.1	4-10
Spelling proficiency (self-rated from 1-10)	7.9	4-10	6.8	4-9
Speaking proficiency (self-rated from 1-10)	9.4	5-10	7.4	5-10
Pronunciation prof. (self-rated from 1-10)	9.2	6-10	6.6	3-9
Reading proficiency (self-rated from 1-10)	9.4	7-10	8.4	6-10
Writing proficiency (self-rated from 1-10)	8.3	5-10	7	4-10
Grammar proficiency (self-rated from 1-10)	7.9	4-10	6.7	4-9

Started hearing the language (age)	0		5.6	1-10
Became fluent in speaking (age)	3.3	2-7	11.7	5-16
Started reading the language (age)	5.2	4-7	7.1	5-10
Became fluent in reading (age)	7.9	5-11	10.7	8-14
L2 intrusion into L1 (accidental from 0-10)	3.3	0-8		
L1 intrusion into L2 (accidental from 0-10)			2.7	0-3
L2 inclusion into L1 (intentional from 0-10)	3.8	0-9		
L1 inclusion into L2 (intentional from 0-10)			2.1	0-5

Appendix D1- English Vocabulary Test- Full stimuli set

Synonyms

Item	**Word**	**Correct**	**FoilA**	**FoilB**	**FoilC**
1	caprice	whim	cattle	brute	lounge
2	baffle	confuse	hide	warp	bully
3	ponderous	unwieldy	useless	supportive	thoughtful
4	banter	chatting	whispering	denial	beating
5	garish	tasteless	spiky	green	bland
6	sequin	bead	stamp	sledge	order
7	loquacious	talkative	broad	roomy	marshy
8	covet	desire	pad	cradle	cave
9	acumen	cleverness	blame	spicy	wealth
10	drench	soak	raise	erase	flatten
11	abide	endure	inhabit	crave	depart
12	vocation	occupation	holiday	pronunciation	vocabulary
13	gulch	crevasse	swallow	shed	dislike
14	cogitate	ponder	achieve	succeed	enquire
15	vexatious	effortful	engaging	horrifying	priceless
16	peril	danger	shiny	delight	shelter
17	feral	savage	hungry	impartial	ugly
18	ludicrous	ridiculous	developed	nasty	certain
19	brisk	energetic	disposable	section	stern
20	truculent	defiant	delicious	juicy	tardy

Antonyms

Item	Word	Correct	FoilA	FoilB	FoilC
1	concerned	uncaring	scarce	misleading	understanding
2	timorous	fearless	forestry	funny	emotive
3	disdain	admire	unload	misfortune	huge
4	acerbic	sweet	itchy	loud	beautiful
5	nonplus	enlighten	subtract	gain	disadvantage
6	surfeit	lack	southern	excess	fake
7	vicious	gentle	slippery	fierce	disobedient
8	saunter	rush	fry	punish	daydream
9	slipshod	careful	difficult	clumsy	footwear
10	umbrage	delight	dungeon	demanding	appeal
11	strenuous	effortless	arduous	smooth	tricky
12	divulge	conceal	purchase	disclose	smuggle
13	loathe	cherish	rejoice	kindle	undress
14	querulous	agreeable	feathered	blatant	squeaky
15	forgo	acquire	precede	journey	disappear
16	conquer	surrender	demand	retain	release
17	hovel	palace	float	cloudy	stairwell
18	adversity	advantage	delay	grudge	persevere
19	alacrity	slowness	annoyance	fog	ingenuity
20	penury	wealth	dispatch	cunning	famine

Appendix D2- Norwegian Vocabulary Test- Full stimuli set

Synonyms

Item	Word	Correct	FoilA	FoilB	FoilC
1	lektyre	lesestoff	leker	hytte	husdyr
2	ufortrøden	uforstyrrelig	uforbederlig	ufokusert	fornøden
3	noksagt	dumrian	ferdigstilt	selvdyrker	påstand
4	lemfeldig	forsiktig	uberegnelig	langsom	frimodig
5	febrilsk	hektisk	illevarslende	tilstrekkelig	varmblodig
6	brudulje	slagsmål	ekteskap	floke	etterligning
7	fjetre	lamme	røpe	legere	finne
8	vankelmodig	ubestemt	nådeløs	mangelfull	hyklersk
9	attrå	begjære	fornærme	avslå	trampe
10	kryste	klemme	brodere	savne	forfølge
11	amper	hissig	skyldig	travel	fyldig
12	smektende	lengtende	spinkel	smakfull	buktende
13	maroder	utmattet	blodtørstig	spenstig	hevngjerrig
14	trettekjær	kranglete	grådig	kresen	svak
15	fadese	tabbe	utside	krig	vegring
16	mulkt	bot	dystert	sveiv	svalt
17	atal	plagsom	sløv	dyktig	hvass
18	vansmekte	lide	gnage	avsky	forgifte
19	sondre	skille	undersøke	forske	vise
20	omkalfatre	endevende	oppfatte	omkomme	omlegge

Antonyms

Item	Word	Correct	FoilA	FoilB	FoilC
1	lapidarisk	pratesyk	usann	kortfattet	fremmed
2	distré	oppmerksom	utakknemlig	motsatt	fordelt
3	sjofel	hyggelig	annerledes	lumpen	skjærende
4	vanvidd	fornuft	ordstrid	viktighet	velklang
5	armod	rikdom	avsporing	elendighet	bopel
6	overflod	fattigdom	omskifte	flom	vrede
7	avertere	skjule	tirre	kunngjøre	forstyrre
8	nennsom	voldsom	sparsom	virksom	strevsom
9	ødsle	spare	hevde	nære	tvile
10	bebreide	berømme	beleire	betvile	betenke
11	uaffisert	påvirket	redigert	offentlig	merkelig
12	besynderlig	alminnelig	snevert	omfattende	anerkjent
13	ublu	rimelig	skjør	freidig	skral
14	hovmod	ydmykhet	angst	avskjed	tilregnelighet
15	anfektelse	visshet	forhindring	åpenbaring	straff
14	petimeter	slask	lekmann	tommestokk	skritt-teller
17	avferdige	godta	avslutte	forhindre	testamentere
18	bifalle	avvise	tilta	snuble	erobre
19	fetere	overse	pine	ernære	flytte
20	nidkjær	slurvete	trassig	selvopptatt	streng

Appendix D3- Auditory Working Memory Test- Full stimuli set

Orto1	Orto2	Condition	Block	Syllables
baarrch teeg nup gook	baarrch teeg nup gook	Same	0	4
baarrch teeg nup gook	baarrch teeg gook nup	Different	0	4
kib dern putch geed	kib dern putch geed	Same	0	4
kib dern putch geed	kib putch dern geed	Different	0	4
maat chen ped kig	maat chen ped kig	Same	0	4
maat chen ped kig	maat ped chen kig	Different	0	4
merd gaarrp tam pib	merd gaarrp tam pib	Same	0	4
merd gaarrp tam pib	merd tam gaarrp pib	Different	0	4
teck kam mitch baan derp	teck kam mitch baan derp	Same	1	5
teck kam mitch baan derp	teck mitch kam baan derp	Different	1	5
choom kerp lork nug gaarrn	choom kerp lork nug gaarrn	Same	1	5
choom kerp lork nug gaarrn	choom kerp nug lork gaarrn	Different	1	5
peem taarrg gab baak chel	peem taarrg gab baak chel	Same	1	5
peem taarrg gab baak chel	peem gab taarrg baak chel	Different	1	5
goot baarrg mern nuck tep	goot baarrg mern nuck tep	Same	1	5
goot baarrg mern nuck tep	goot baarrg nuck mern tep	Different	1	5
loog jaal didge kerm meb	loog jaal didge kerm meb	Same	1	5
loog jaal didge kerm meb	loog jaal kerm didge meb	Different	1	5
keech jaarrn mep terg bick	keech jaarrn mep terg bick	Same	1	5
keech jaarrn mep terg bick	keech mep jaarrn terg bick	Different	1	5
koom torrd maadge jup gick	koom torrd maadge jup gick	Same	1	5
koom torrd maadge jup gick	koom torrd jup maadge gick	Different	1	5
laad tudge jick norb gaarrm	laad tudge jick norb gaarrm	Same	1	5
laad tudge jick norb gaarrm	laad jick tudge norb gaarrm	Different	1	5
bordge chud neeg dack keb larm	bordge chud neeg dack keb larm	Same	2	6
bordge chud neeg dack keb larm	bordge neeg chud dack keb larm	Different	2	6
terdge joop leck norrg chim peeb	terdge joop leck norrg chim peeb	Same	2	6
terdge joop leck norrg chim peeb	terdge joop norrg leck chim peeb	Different	2	6
paarrn mab dorge naag cheem jit	paarrn mab dorge naag cheem jit	Same	2	6
paarrn mab dorge naag cheem jit	paarrn mab dorge cheem naag jit	Different	2	6
kern boodge tud lig pab dorrt	kern boodge tud lig pab dorrt	Same	2	6
kern boodge tud lig pab dorrt	kern tud boodge lig pab dorrt	Different	2	6
tidge mup chen gerb noog deet	tidge mup chen gerb noog deet	Same	2	6
tidge mup chen gerb noog deet	tidge mup gerb chen noog deet	Different	2	6
torrm pag ieck derb kaal bup	torrm pag ieck derb kaal bup	Same	2	6
torrm pag ieck derb kaal bup	torrm ieck pag derb kaal bup	Different	2	6
korrp teeb nool jaarrk pim gerch	korrp teeb nool jaarrk pim gerch	Same	2	6
korrp teeb nool jaarrk pim gerch	korrp teeb jaarrk nool pim gerch	Different	2	6
padge naarrp maan chut goob ged	padge naarrp maan chut goob ged	Same	2	6
padge naarrp maan chut goob ged	padge naarrp maan goob chut ged	Different	2	6
toock chad lidge jerg dop naarrt gub	toock chad lidge jerg dop naarrt gub	Same	3	7
toock chad lidge jerg dop naarrt gub	toock chad lidge dop jerg naarrt gub	Different	3	7
kaarrk nertch morrd cham bool lub teep	kaarrk nertch morrd cham bool lub teep	Same	3	7
kaarrk nertch morrd cham bool lub teep	kaarrk nertch morrd cham lub bool teep	Different	3	7
chaarrn nig kaam jeel gadge lerb mun	chaarrn nig kaam jeel gadge lerb mun	Same	3	7
chaarrn nig kaam jeel gadge lerb mun	chaarrn kaam nig jeel gadge lerb mun	Different	3	7
taab gan daarrch chool juck norrd pem	taab gan daarrch chool juck norrd pem	Same	3	7
taab gan daarrch chool juck norrd pem	taab gan chool daarrch juck norrd pem	Different	3	7
gel nerg lud paack meetch doob jat	gel nerg lud paack meetch doob jat	Same	3	7
gel nerg lud paack meetch doob jat	gel nerg lud paack doob meetch jat	Different	3	7
leem kug chordge jert ked daarrp gaack	leem kug chordge jert ked daarrp gaack	Same	3	7
leem kug chordge jert ked daarrp gaack	leem kug chordge ked jert daarrp gaack	Different	3	7
chig nam peb gaap jooch laarrt teed	chig nam peb gaap jooch laarrt teed	Same	3	7
chig nam peb gaap jooch laarrt teed	chig peb nam gaap jooch laarrt teed	Different	3	7
jaarrm neb gerp chorrg mal tooch larn	jaarrm neb gerp chorrg mal tooch larn	Same	3	7
jaarrm neb gerp chorrg mal tooch larn	jaarrm neb gerp mal chorrg tooch larn	Different	3	7
choom kerp lork nug gaarrn	choom kerp lork gaarrn nug	Different	1	5
koom torrd maadge jup gick	koom torrd maadge gick jup	Different	1	5
korrp teeb nool jaarrk pim gerch	korrp teeb nool jaarrk gerch pim	Different	2	6
leem kug chordge jert ked daarrp gaack	leem kug chordge jert ked gaack daarrp	Different	3	7

Appendix D4- LexTALE Test- Full stimuli set

(Retrieved from lextale.com//downloads/ExperimenterInstructionsEnglish.pdf 04 May 2020)

• Third column: word status; 0=nonword, 1=word.

Item number	Item	
0	platery	0
0	denial	1
0	generic	1
1	mensible	0
2	scornful	1
3	stoutly	1
4	ablaze	1
5	kermshaw	0
6	moonlit	1
7	lofty	1
8	hurricane	1
9	flaw	1
10	alberation	0
11	unkempt	1
12	breeding	1
13	festivity	1
14	screech	1
15	savoury	1
16	plaudate	0
17	shin	1
18	fluid	1
19	spaunch	0
20	allied	1
21	slain	1
22	recipient	1
23	exprate	0
24	eloquence	1
25	cleanliness	1
26	dispatch	1
27	rebondicate	0
28	ingenious	1
29	bewitch	1
30	skave	0

Item number	Item	
31	plaintively	1
32	kilp	0
33	interfate	0
34	hasty	1
35	lengthy	1
36	fray	1
37	crumper	0
38	upkeep	1
39	majestic	1
40	magrity	0
41	nourishment	1
42	abergy	0
43	proom	0
44	turmoil	1
45	carbohydrate	1
46	scholar	1
47	turtle	1
48	fellick	0
49	destription	0
50	cylinder	1
51	censorship	1
52	celestial	1
53	rascal	1
54	purrage	0
55	pulsh	0
56	muddy	1
57	quirty	0
58	pudour	0
59	listless	1
60	wrought	1

Appendix E- Language Proficiency Tests- Results

subject_nr	aud_WM_acc	Vocab_No_acc	Vocab_en_acc	lextale_en_acc
1	63,3	10	35	92,2
2	56,7	50	30	87,3
3	50	42,5	27,5	84,1
4	63,3	30	27,5	76,2
5	56,7	30	37,5	73
6	66,7	42,5	40	88,9
7	66,7	20	57,5	92,1
8	80	30	57,5	88,9
9	70	45	60	88,9
10	76,7	35	12,5	71,4
11	53,3	25	20	63,5
12	66,7	35	60	87,3
13	73,3	20	30	85,7
14	80	52,5	27,5	92,1
15	73,3	30	15	73
16	80	50	20	69,8
17	70	57,5	27,5	68,3
18	66	57,5	30	71,4
19	70	40	15	66,7
20	66,7	20	15	79,4
21	76	45	40	92,1
22	53,3	22,5	40	93,7
23	60	12,5	27,5	82,5
24	80	22,5	40	90,5
25	46,7	30	20	77,8
26	56,7	15	20	77,8
27	53,3	47,5	40	85,7
28	60	27,5	27,5	87,3
29	66,7	15	32,5	87,3
31	63,3	17,5	7,5	79,4
32	86,7	40	40	84,1
34	66,7	32,5	37,5	87,3
35	53,3	32,5	30	92,1
	65,8	**32,8**	**31,7**	**82,4**

Appendix F1- Visual Memory Test- Word stimuli set

Counter	Scene	Initial noncognate	Object number	Object tested	Cognates	Noncognates
1	nursery	blackboard	2		pie	custard
			3	x	rattle	kitten
			4		elf	puppet
2	babyroom	toy train	2		mattress	cradle
			3	x	pony	donkey
			4		teepee	fortress
3	bedroom1	big bed	2	x	pyjamas	robe
			3		radio	tissue box
			4		shoes	slippers
4	bedroom2	double bed	2		clock	mirror
			3	x	stool	rocker
			4		jumpsuit	trousers
5	window_sill	child's drawing on the wall	2		compass	sundial
			3	x	brooch	bracelet
			4		beetle	cricket
6	dresser	necklace	2	x	perfume	moisturiser
			3		scarf	handkerchief
			4		yarn	ribbons
7	white_bathroom	bathrobe	2	x	bathmat	rug
			3		toilet paper	towels
			4		shampoo	conditioner
8	bathroom_counter	hairbush	2		cream	cologne
			3	x	thermometer	toothpaste
			4		soap	sponge
9	kitchen_counter	measuring jug	2		pear	orange
			3	x	aubergine	zucchini
			4		eggs	cheese
10	kitchen_with_spice_rack	spice rack	2		marmelade	mustard
			3	x	milk	detergent
			4		plums	gooseberries

			2		olive oil	vinegar
11	kitchen_rack	scales	3	x	coffee	flour
			4		rice	lentils
12	white_kitchen	coffee maker	2	x	blender	grinder (meat)
			3		salad	cabbage
			4		pasta	cereal
13	old_kitchen_with_pie	water melon	2		bread	buns (3 stk)
			3	x	clementine	lemon
			4		broccoli	carrots
14	kitchen_under_construction	dishwasher	2	x	tomato	pepper (red)
			3		plastic (protective)	bag (tarp)
			4		oven	cooker
15	green_livingroom	side table	2		bench (benk)	chair
			3	x	beanbag	cushion
			4		window	curtains
16	buddha_livingroom	buddha statue	2		piano	keyboard
			3	x	plant	flowers
			4		sword	spear
17	red_livingroom	Christmas tree	2		ballerina	doll
			3	x	racket	bat
			4		violin	fiddle
18	fancy_livingroom	staircase	2		portrait	picture
			3	x	sofa	couch
			4		puff	dogbed
19	livingroom_table	table ornament	2	x	shirt	sweater
			3		passport	diary
			4		calculator	remote
20	livingroom_with_safe	two large speakers	2		cabinet	coat rack
			3	x	tambourine	drum
			4		safe	painting
21	fireplace	fire extinguisher	2	x	angel	santa
			3		bust	hourglass
			4		apple	pumpkin

22	bar	menu	2		wine	cutlery
			3	x	icecream	sundae
			4		ketchup	herbs
23	wood_diningroom	crystal goblet	2	x	globe	candle
			3		puzzle	chessboard
			4		bible	dictionary
24	table_for_two	chair with a shawl on it	2	x	glass	jug
			3		wine list	tray
			4		menu	advertisement
25	white_table	deck of cards	2		lighter	matches
			3	x	cupcake	pastry
			4		salt	cinnamon
26	corner_table	reed diffuser	2	x	telephone	wallet
			3		credit card	envelope
			4		sandals	wellies
27	coffee_table	model of the eiffel tower	2		mango	raspberries
			3	x	banana	corncob
			4		headset	earmuffs
28	hallway	open door	2		rocket	candlestick
			3	x	radio	briefcase
			4		bowl	dish
29	wall_hooks	keys	2	x	hat	basket
			3		shoehorn	umbrella
			4		cap	beanie
30	clothes_rack	mannequin	2	x	skirt	dress
			3		moccasines	skates
			4		(spool of) thread	zipper
31	storeroom	extension cord	2		tulip	daffodil
			3	x	printer	toolbox
			4		helmet	hardhat
32	storage	dustpan	2	x	drill	hairdryer
			3		oil can	paint can
			4		skis	pipe

33	blue_house	swing	2		basketball	steering wheel
			3	x	megaphone	traffic cone
			4		taxi	hearse
34	green_basement	sewing machine on a table	2		freezer	dryer
			3	x	chest	suitcase
			4		mop	broom
35	red_basement	stack of shelves	2	x	wheel	tyre
			3		potatoes (sack)	onions (sack)
			4		case	cooler
36	wine_cellar	wine barrel	2		television	wood burner
			3	x	sleigh	toboggan
			4		rat	ferret
37	balcony	yoga mat	2		garland	festoon
			3	x	recliner	planter
			4		stone	rock
38	big_balcony	telescope	2	x	statue	firepit
			3		fountain	pool
			4		grill	stove
39	outdoor_patio	bird house	2	x	(box of) chocolate(s)	biscuits
			3		camera	binoculars
			4		bee	spider
40	patio_table	baby car seat	2	x	melon	pineapple
			3		hammer	scissors
			4		mouse	squirrel
41	backyard_picnic	picnic blanket	2	x	pheasant	peacock
			3		skateboard	tricycle
			4		bush	shrub
42	backyard_rabbit	trailer	2		trampoline	pond
			3	x	pavilion	Cottage
			4		hare	Rabbit
43	garden_steps	butterfly	2	x	Pot	bucket
			3		Snail	slug
			4		Frog	toad

44	garden_shed	lawnmower	2	x	Rake	hoe
			3		Saw	shovel
			4		axe	hatchet
45	garden_with_well	well	2	x	Goose	duck
			3		cat	Dog
			4		(roll of) wire	(roll of) rope
46	garage	folding chair	2	x	Battery	Typewriter
			3		Scooter	bike
			4		Canoe	Ladder
47	snowy_street	large inflatable snowman	2	x	tractor	car
			3		snowplow	truck
			4		crow	seagull
48	street_with_garbage	fountain	2	x	Dove	pigeon
			3		mink	weasel
			4		screws	padlock
49	street_with_graffiti	telephone booth	2	x	Graffiti	posters (plakat)
			3		Paper	Dirt
			4		Bus	lorry
50	picnic_table	burger	2	x	thermos	bottle
			3		frisbee	Plate
			4		walnuts	acorns
51	playground	climbing frame	2	x	tunnel	slide
			3		tree	pole
			4		boat	train
52	wedding_reception	a row of chairs	2	x	cake	trifle
			3		bouquet	bow
			4		quartet	performer
53	birthday_party	playmat	2		clown	magician
			3	x	package	present
			4		balloons	streamer
54	reception_table	vending machine	2		calendar	screen
			3	x	vase	jar
			4		brochure	pamphlet

55	hotel_room	palm tree	2		eye-shadow	purse
			3	x	belt	strap
			4		chapel	lighthouse
56	construction_site	traffic sign	2		bulldozer	excavator
			3	x	sand	gravel
			4		barrier	gate
57	executive_office	stack of files	2		laptop	folder
			3	x	cup	mug
			4		(pack of) cigarettes	(pack of) chewing gum
58	student_office	binders	2	x	yoghurt	seeds
			3		postcard	letter
			4		CD	record
59	chemistry_lab	microscope	2	x	goggles	glasses
			3		beaker (beger)	test tube
			4		skeleton	chart
60	conference_room	vacuum cleaner	2		pens	pencils
			3	x	pizza	newspaper
			4		marker	chalk
61	throne_room	tapestry	2	x	throne	armchair
			3		shield	crest
			4		rifle	Gun
62	waiting_room	chairs	2	x	carton	cage
			3		magazine	clipboard
			4		Diploma	(poster of a) Brain
63	police_room	cash	2		tape	recorder
			3	x	mask	wig
			4		notebook	map
64	gym	punching bag	2	x	mat	sweatshirt
			3		springboard	scales
			4		radiator	weights
65	gym_locker	yellow warning sign	2		t-shirt	sweatpants
			3	x	socks	gloves
			4		duffelbag	kitbag

66	barnyard	dog house	2		goat	ram
			3	x	hen	rooster
			4		cow	pig
67	museum	signs	2	x	panther	cheetah
			3		manuscript	scroll
			4		fossil	pottery
68	museum_display	horse statue	2	x	knife	dagger
			3		egg	gem (size of egg)
			4		skull	sea urchin
69	scandinavian_loft	poster	2		coal	kindling
			3	x	baggage	backpack
			4		golf clubs	walking sticks
70	country_house	rose bush	2		lantern (on wall)	torch (on wall)
			3	x	llama	sheep
			4		calf	stag
71	mansion	pedestal	2		rose	daisy
			3	x	chaise longue	hammock
			4		urn	pitcher
72	white_beach	sunbed	2	x	crab	lobster
			3		surfboard	shark
			4		sailboat	freighter
73	beach_house	anchor	2		net	tripod
			3	x	cocktail	beer
			4		bikini	trunks
74	church	donations box	2	x	cactus	fern
			3		cross	wreath
			4		lamp	fan
75	wheat_field	some sunflowers	2		windmill	scarecrow
			3	x	oak	maple
			4		silo	steeple
76	barn	boots	2	x	raincoat	overalls
			3		saddle	blanket
			4		straw	sawdust

77	cliffs	caravan	2	x	tent	lounger
			3		fire	puddle
			4		albatros	eagle
78	jungle	giraffe	2	x	gorilla	monkey
			3		fish	turtle
			4		leopard	lion
79	subway	bag of chips	2		shawl	leash
			3	x	crutch	cane
			4		poodle	puppy
80	stage	Guitar	2		podium	music stand
			3	x	popcorn	candy
			4		singer	conductor

Appendix F2- Visual Memory Test- Sentence stimuli sets

Target scene sentences

Scene	Sentence (cognate/noncognate)
nursery	This is a nursery: there is a blackboard, a pie/some custard, a rattle/kitten, and an elf/puppet on the floor.
babyroom	This is a babyroom: there is a toy train, a mattress/cradle, a pony/donkey, and a teepee/fortress.
bedroom1	This is a bedroom: there is a big bed, a set of pyjamas/a robe hanging on the wall, a radio/tissue box, and a pair of shoes/slippers
bedroom2	This is a large bedroom: there is a bed, a clock/mirror, on the wall, a stool/rocker, and a jumpsuit/trousers on the floor.
window_sill	This is a windowsill: there is a child's drawing on the wall, and there is a compass/sundial, a brooch/bracelet, and a beetle/cricket by the window.
dresser	This is a dresser: there is a necklace stand, a perfume bottle/moisturizer bottle, a scarf/handkerchief in a drawer, and a basket with yarn/ribbons on the floor.
white_bathroom	This is a white bathroom: there is a white bathrobe, a bathmat/rug, some toilet paper/towels, and a bottle of shampoo/conditioner.
bathroom_counter	This is a bathroom counter: there is a hairbrush, some cream/cologne, a thermometer/some toothpaste, and a bar of soap/a sponge.
kitchen_counter	This is a kitchen counter: there is a measuring jug, a pear/orange, an aubergine/zucchini, and some eggs/cheese
kitchen_with_spice_rack	This is an old kitchen: there is a spice rack, some marmalade/mustard, a bottle of milk/detergent, and some plums/gooseberries.
kitchen_rack	This is a kitchen rack: there are some scales, a bottle with olive oil/vinegar, a bag of coffee/flour, and a bag of rice/lentils on the floor.
white_kitchen	This is a large white kitchen: there is a coffee maker, a blender/grinder, a salad/some cabbage, and a box of pasta/cereal.
old_kitchen_with_pie	This is an old kitchen: there is a watermelon on one counter, and on the other counter, there is some bread/some buns, a clementine/lemon, and some broccoli/carrots
kitchen_under_construction	This is a kitchen under construction: there is a dishwasher, a tomato/pepper on the floor, a roll of plastic/a bag, and an oven/cooker in the corner.
green_livingroom	This is a green living room: there is a side table, a bench/chair, a beanbag/cushion, and a large window/curtains.
buddha_livingroom	This is a living room: there is a buddha statue, a piano/keyboard, a plant/some flowers, and a sword/spear above the fireplace.
red_livingroom	This is a large living room: there is a Christmas tree, a ballerina/doll, a racket/bat, and a violin/fiddle.
fancy_livingroom	This is a fancy living room: there is a staircase, a portrait/picture on the wall, a sofa/couch, and a puff/dog bed.
livingroom_table	This is a living room table: there is a table ornament, a folded shirt/folded sweater, a passport/diary, and a calculator/remote.
livingroom_with_safe	This is a beige living room: there are two large speakers, a cabinet/coat rack, a tambourine/drum, and a safe/painting on the wall.
fireplace	This is a room with a fireplace: there is a fire extinguisher on the floor, and there is an angel/Santa, a bust/hourglass, and an apple/pumpkin on the mantlepiece,
bar	This is a diner: there is a menu on the table, a bottle of wine/some cutlery, an ice-cream/sundae, and a bottle of ketchup/some herbs on the seat.
wood_diningroom	This is an old dining room table: there is a crystal goblet, a globe/candle, a puzzle/chessboard, and a bible/dictionary.
table_for_two	This is a table in an outdoor restaurant: there is a chair with a shawl on it, on the table there is a glass/jug and a wine list/serving tray, and there is a board with a menu/advertisement against the back wall.
white_table	This is a white dining room table: there is a deck of cards, a lighter/a box of matches, a cupcake/pastry, and salt/cinnamon.
corner_table	This is a corner table: there is a scent diffuser, a telephone/wallet, a credit card/envelope, and some sandals/wellies on the floor.
coffee_table	This is a small coffee table: there is a model of the Eiffel tower, a mango/some raspberries, a banana/corncob, and a headset/earmuffs.
hallway	This is a hallway: there is an open door, a model of a rocket/a candlestick, an old radio/briefcase, and a bowl/dish.

wall_hooks	This is a coat rack: there is a key, a hat/basket, a shoehorn/umbrella, and a cap/beanie.
clothes_rack	This is a clothes rack: there is a mannequin next to the rack, and there is a skirt/dress, a pair of moccasins/skates, and a spool of thread/zipper on the floor
storeroom	This is a storeroom counter: there is an extension cord, a bouquet of tulips/daffodils, a printer/toolbox, and a helmet/hardhat.
storage	This is a dark storage room: there is a dustpan on the floor, a drill/hairdryer, an oil can/paint can, and some skis/pipes.
blue_house	This is a driveway next to a blue house: there is a swing, a basketball/steering wheel, a megaphone/traffic cone, and a black taxi/hearse.
green_basement	This is a green basement: there is a sewing machine on the table, a freezer/dryer, a chest/suitcase, and a mop/broom.
red_basement	This is a red basement: there is a stack of shelves, a wheel/tyre, a sack of potatoes/onions, and a case/cooler.
wine_cellar	This is a wine cellar: there is a wine barrel, a television/wood burner, a sleigh/toboggan, and a rat/ferret.
balcony	This is a balcony: there is a yoga mat, a garland/a festoon, a recliner/planter, and a stone/rock.
big_balcony	This is a big balcony: there is a telescope, a statue/firepit, a fountain/pool, and a grill/ stove.
outdoor_patio	This is an outdoor patio: there is a bird house, and on the table, there is a box of chocolates/biscuits, a camera/binoculars, and a bee/spider.
patio_table	This is a patio table: there is a baby car seat, a melon/pineapple, a hammer/scissors, and a mouse/squirrel.
backyard_picnic	This is a backyard: there is a picnic blanket, a pheasant/peacock, a skateboard/tricycle, and a bush/shrub in a pot to be planted.
backyard_rabbit	This is a large backyard: there is a trailer, a trampoline/pond, a pavilion/cottage, and a hare/rabbit in the grass
garden_steps	These are steps in a garden: there is a butterfly, a pot/bucket, a snail/slug, and a frog/toad
garden_shed	This is a garden shed: there is a lawnmower, a rake/hoe, saw/shovel, and an axe/hatchet.
garden_with_well	This is a garden: there is a well, a goose/duck, a sleeping cat/dog, and a roll of wire/rope.
garage	This is a garage: there is a foldable chair outside, and inside there is battery/typewriter, a scooter/bike, and a canoe/ladder
snowy_street	This is a street covered in snow: there is a large inflatable snowman, a tractor/car, a snowplough/truck, and a crow/seagull sitting on a railing.
street_with_garbage	This is a street by a park: there is a fountain on the wall, and on the ground, there is a dove/pigeon, a mink/weasel, and a pile of screws/a padlock.
street_with_graffiti	this is an empty street: there is an old telephone booth, some graffiti/posters on the wall, some paper/a pile of dirt on the ground, and a bus/lorry in the distance
picnic_table	This is a picnic table: there are some burgers, a thermos/bottle, a frisbee/plate, and some walnuts/acorns
playground	This is an empty playground: there is a climbing structure, a tunnel/slide, a tree/pole, and a boat/train.
wedding_reception	This is a wedding reception: there are some seats against the wall, a table with a cake/trifle, a chair with a bouquet/bow, and there is a quartet/performer in the back of the room.
birthday_party	This is a birthday party: there is a playmat on the floor, a clown/magician, a package/present, and balloons/streamer by the stairs.
reception_table	This is a reception: there is a vending machine in the corner, and on the counter, there is a calendar/monitor, a vase/jar, and some brochures/pamphlets.
hotel_room	This is a hotel room: there is a palm tree by the window, and on the bed, there is an eye-shadow set/a purse, a belt/strap, and a picture of a chapel/lighthouse on the wall
construction_site	This is a construction site: there is a traffic sign, a bulldozer/excavator, a pile of sand/gravel, and a barrier/gate.
executive_office	This is an executive office: on the table, there is a stack of files, a laptop/folder, a cup/mug, and a pack of cigarettes/lozenges.
student_office	This is a student's office: on the table, there are some binders, a bowl of yoghurt/bowl of seeds, a postcard/letter, and a cd/record.

chemistry_lab	This is a chemistry lab: there is a microscope, some goggles/glasses, a beaker/test tube, and a poster of a skeleton/chart.
conference_room	This is a conference room: there is a vacuum cleaner by the table, and on the table, there are some pens/pencils, a pizza box/newspaper, and a box of markers/chalk.
throne_room	This is a throne room: there is a tapestry on the wall, and there is a throne/armchair, a shield/crest, and a rifle/gun on the floor.
waiting_room	This is a waiting room: there is a shopping trolly, a carton/cage on the floor, a magazine/clipboard, and on the wall, there is a diploma/poster of a brain.
police_room	This is a police room: there is some cash, a roll of tape/recorder, a mask/wig, and a notebook/map.
gym	This is a large empty gym: there is a punching bag hanging from the ceiling, and on the floor, there is a mat/sweatshirt, a springboard/scale, and a radiator/some weights.
gym_locker	This is a gym locker room: there is a yellow warning sign on the floor, and there is a t-shirt/sweatpants hanging on a locker door, some socks/gloves on the floor, and a duffel bag/kitbag.
barnyard	This is a farmhouse: there is a red doghouse by the side of the house, and in the grass there is a goat/ram, a hen/rooster, and a cow/pig.
museum	This is a room in a museum: there is a stand with a sign, a panther/cheetah, a manuscript/scroll, and a fossil/some pottery.
museum_display	This is museum cabinet: there is a statue of a horse, a knife/dagger, an egg/gem, and a skull/sea urchin
scandinavian_loft	This is a loft: there is a rolling stones poster, some coal/kindling, baggage/a backpack, and golf clubs/walking sticks.
country_house	This is a country house: there is a rose bush by the house, a lantern/torch on the wall, and in the grass, there is a llama/sheep and a calf/stag.
mansion	This is a room in a mansion: there is a pedestal, a picture of a rose/daisy, a chaise longue/hammock, and an urn/pitcher.
white_beach	This is a sunny beach: there is a sun bed, a crab/lobster, a surfboard/shark, and a sailboat/freighter in the distance
beach_house	This is a beach house: there is an anchor on the wall, a net/tripod by the wall, a cocktail/beer and a bikini/trunks on the floor.
church	This is a church hallway, there is a donations box, a cactus/fern on the table, a cross/wreath on the wall, and a lamp/fan.
wheat_field	This is a wheat field: there is a sunflower, a windmill/scarecrow, an oak/maple on the horizon, and a silo/steeple.
barn	This is a barn: there are some boots, a raincoat/overalls, a saddle/some blankets, and some straw/sawdust.
cliffs	This is a field by the cliffs: there is a caravan, a tent/lounger, a fire/puddle, and an albatross/eagle flying overhead.
jungle	This is a jungle: there is a giraffe, a gorilla/monkey, a fish/turtle in the water, and a leopard/lion.
subway	This is a subway train: there is a bag of chips on the seat, a shawl/leash, a crutch/cane, and a poodle/puppy sitting down.
stage	This is a stage: there is a guitar, a podium/music stand, a bowl of popcorn/candy, and a singer/conductor.

Filler scene sentences

This is an alleyway: there is shopping trolly, a fire hydrant, a cash machine, and a park meter
This is a changing room: there are some shopping bags, a pair of high heels, and a stack of folded jeans, and a ballerina poster on the wall
This is a photography studio: there are some huge lamps, a long roll of white paper, a tall seat, and an iPad on the floor
This is a hillside: there is a boy running and looking up at a kite
This is a square in an Italian city: there is a mailbox and an olive tree in the middle, and there are two children running towards an archway
This is a swimming pool: there are two children in the water and there is a high seat with a lifeguard and inflatable ball on the side of the pool
This is a train station: there is a train coming in and one man waiting close to a ticket machine.
This is a laundry room: there are some cupboards, a pile of clothes, and a washing machine with an iron on top
This is a swimming pool: there is a woman stretching, a springboard and a rubber ring in the water.
This is an old wall: there is path running along side the wall, and there is a strip of grass and a small bike stand.
This is a room in an abandoned house: there is chimney, a warning sign on the wall, and a lot of rubble.
This is a loft room with wooden floors and bare brick walls: there is a hanging chair, a shark decoration on the wall, and a xmas stocking by the window
This is a large field: there is a wheelbarrow full of grass clippings, a cricket set, and there a two horses grazing in the distance
This is a garden: it is very green, there is a swing, a see saw, a garden hose, and some stone houses in the background.
This is a castle hallway: there is a suit of armour, a flag on the wall and a postcard stand.
This is a bachelor pad: there is a dartboard, a foldable table with a box of takeaway food, a video game console, and a poster of Marilyn Monroe by the window
This is a bedroom under construction: there is a screen, a roll of bubble wrap, an ipod on the floor, and a recycling bin
This is a home gym: there is a treadmill by the window, and there on the pall there is a poster of a football player and a hanging paper star
This is a desert: there are two camels with a man on one of them and there is a small village with huts and palm trees in the distance.
This is a lake surrounded by cliffs: there is a wooden cabin by the water and there is a person sitting on one of the rocks holding a fishing rod.

Appendix G- Maths Questions

1	Answer		Answer		Answer		Answer
18+33=		11+34=		32+8=		520+73=	
10+4=		66+2=		3+59=		11+44=	
3+19=		102+87=		95+8=		69+2=	
55+2=		4+13=		92+7=		12+47=	
72+3=		21+1=		9+133=		7+143=	
7+33=		99+6=		18+31=		12+83=	
2+81=		2+67=		96+23=		26+11=	
68+1=		33+29=		12+7=		100+34=	
12+34=		7+6=		4+14=		22+71=	
14+9=		19+17=		32+7=		81+6=	
2+16=		3+3=		2+39=		12+33=	
42+1=		20+6=		155+3=		6+24=	
5+9=		65+97=		902+7=		72+8=	
35+8=		23+5=		17+33=		14+19=	
6+19=		35+2=		4+83=		8+6=	
3+99=		2+197=		54+1=		436+2=	
54+2=		71+6=		199+3=		44+1=	
33+5=		28+13=		4+49=		6+87=	
9+12=		10+45=		2+87=		3+63=	
18+7=		52+73=		22+8=		88+9=	
8+2=		44+8=		34+6=		11+54=	
20+2=		210+67=		22+5=		3+258=	
32+9=		42+6=		9+36=		16+41=	
7+42=		3+55=		19+3=		10+22=	
2+33=		8+6=		14+44=		98+71=	

6+99=	49+2=	72+6=	5+5=
12+91=	4+10=	7+43=	7+9=
64+4=	37+25=	3+727=	8+44=
78+4=	34+6=	76+11=	1+156=
249+4=	60+45=	8+8=	3+21=
102+3=	33+3=	14+41=	100+54=
22+9	1+78=	4+18=	6+6=
765+1	86+6=	2+69=	53+66=
72+91	3+80=	854+1=	85+4=
12+3	54+7=	763+3=	7+11=
56+9	43+70=	3+67=	13+91=
32+32	38+6=	4+69=	2+44=
55+63	99+30=	39+8=	6+68=
88+1=	6+9=	5+65=	1+72=
76+9=	4+87=	48+3=	23+5=
38+3=	21+1=	9+133=	17+34=
87+33=	9+6=	28+31=	82+8=
2+81=	5+91=	9+73=	56+12=
8+194=	13+9=	32+6=	612+4=
2+37=	89+6=	75+15=	82+41=
547+9=	19+17=	32+27=	27+26=
19+45=	3+83=	24+9=	17+43=
4+188=	206+6=	75+1=	61+4=
5+89=	73+91	30+62=	4+227=
68+62=	16+39=	87+352=	8+82=
10+66=	23+123=	44+83=	51+40=
3+86=	20+87=	65+13=	4+52=

56+89=	3+92=	44+37=	45+10=
19+3=	1+7643=	203+90=	52+27=
32+31=	8+84=	111+76=	54+4=
14+65=	47+33=	65+45=	15+93=
82+37=	188+2=	23+57=	72+98=
101+95=	50+82=	81+83=	44+76=