



UNIVERSITETET I AGDER

# **Assessing Bilingual Language Proficiency through Self-ratings**

BEATRICE ZITONG URLAND

## **SUPERVISOR**

Linda Ruth Wheeldon  
Allison Louise Wetterlin

## **University of Agder, 2019**

Faculty of Humanities and Education  
Department of Foreign Languages and Translation



## Table of Contents

<b>Introduction .....</b>	<b>1</b>
<b>Bilingual language processing.....</b>	<b>2</b>
<b>The BIA Model.....</b>	<b>2</b>
Orthographic neighborhood effects in bilingual word recognition .....	3
Cognate facilitation and the cascaded activation model .....	6
Phonological overlap between L1 and L2 .....	8
<b>The Revised Hierarchical Model (RHM).....</b>	<b>10</b>
L1 activation during L2 lexical processing.....	12
<b>The Inhibitory Control Model .....</b>	<b>15</b>
Lexical access in bilingual speech production.....	15
<b>Effect of L2 production on language proficiency .....</b>	<b>19</b>
<b>Assessing accent proficiency in bilingualism .....</b>	<b>19</b>
<b>Is L2 verbal fluency an accurate indicator of language proficiency? .....</b>	<b>23</b>
<b>Self-rated language profile .....</b>	<b>26</b>
<b>Self-Assessment of Linguistic Skills .....</b>	<b>26</b>
<b>LEAP-Q.....</b>	<b>28</b>
<b>Norwegian English Language Comparison.....</b>	<b>32</b>
<b>Cognates and homographs .....</b>	<b>33</b>
<b>Phonetic Inventory: Vowels.....</b>	<b>33</b>
<b>Phonetic Inventory: Consonants.....</b>	<b>34</b>
<b>Morphology and Syntax .....</b>	<b>35</b>
<b>Subject-verb agreement .....</b>	<b>37</b>
<b>Prosodic features .....</b>	<b>37</b>

<b>Motivation and predictions .....</b>	<b>39</b>
<b>Methods.....</b>	<b>41</b>
<b>Participants .....</b>	<b>41</b>
<b>Materials and procedure .....</b>	<b>41</b>
<b>Description of battery of behavioral tasks .....</b>	<b>43</b>
<b>Results .....</b>	<b>47</b>
<b>Factors analysis .....</b>	<b>49</b>
<b>Correlations between behavioral and self-reported measures.....</b>	<b>52</b>
<b>Discussion .....</b>	<b>54</b>
<b>Conclusion.....</b>	<b>60</b>
<b>References .....</b>	<b>62</b>
<b>Appendix A.....</b>	<b>66</b>
<b>Appendix B.....</b>	<b>76</b>
<b>Appendix C.....</b>	<b>79</b>

## **Introduction**

The current study investigates bilingual language processing in Norwegian-English bilinguals, and examines the relationship between self-rated language proficiency and standardized measures. This study also aims to investigate the role of self-rated accent proficiency, and whether it can relate to other aspects of second language (L2) proficiency (i.e. reading comprehension, vocabulary, spelling, phoneme production, and grammar). Evidence from previous studies (e.g., Delgado, Guerrero, Goggin, & Ellis, 1999; Marian, Blumenfeld & Kaushanskaya, 2007) suggested that bilinguals' self-ratings of their language skills provided a reliable account of their overall language proficiency. The goal of the current research is to further examine the reliability of self-assessed language proficiency, especially by including self-assessed accent proficiency as a language variable. This study has adopted and extended the Language Experience and Proficiency Questionnaire (LEAP-Q), created by Marian, Blumenfeld, and Kaushanskaya (2007). Aside from modifications to the questionnaire, the current study has also included more items for the standardized tests, in order to reflect a more comprehensive measurement of the bilinguals' overall L2 performance.

While Norwegian and English are both Germanic languages with similarities in areas such as orthography, phonology, and grammar, there are also quite a few distinctions that set the two languages apart. Unlike English, there is no standard variety (i.e., Standard American English or Received Pronunciation in Britain) in Norwegian. Furthermore, regional dialects in Norway can vary quite drastically depending on which part of the country the speakers come from. The current study examines the bilingual language proficiency of high-level Norwegian-English speakers through self-assessed and standardized language measurements. This study also investigates the different factors that could influence the bilinguals' language proficiency (i.e. age of acquisition, extent of L2 exposure, language attitude and preference, etc.), with a special emphasis on the relationship between self-rated accent proficiency and other aspects of their L2 proficiency. There are several bilingual language processing models (i.e. the BIA model, the RHM, and the Inhibitory Control Model) explaining the workings of the bilingual mind, especially in terms of the selection process between the bilinguals' first/native language (L1) and L2.

In this paper, I will first review the role of language proficiency in current models of bilingual language processing, along with key studies to provide evidence for each language model (e.g., Van Heuven, Dijkstra & Grainger, 1998; Kroll & Stewart, 1994; Costa, Caramazza, &

Sebastian-Galles, 2000; Costa & Santesteban, 2004; Sunderman & Kroll, 2006; Green, 1986). I will then discuss the role of self-rated language proficiency, and why it has been deemed as an effective tool for evaluating bilingual language proficiency (e.g., Delgado et al., 1999; Marian et al., 2007). Lastly, I will review the role of L2 speech production (both in terms of accent and fluency) and its relationship to actual linguistic performances (e.g., Chakraborty, Domsch, and Gonzales, 2011; De Jong, Groenhout, Schoonen, & Hulstijn, 2013). After the literature review, I will discuss the main structural differences between Norwegian and English (i.e., phonetic inventory, morphology and syntax, sentence structure, and prosodic features), and how they informed the construction of our language experiments. Finally, I will discuss methods for assessing proficiency, before moving on to results and discussion.

## **Bilingual language processing**

### **The BIA Model**

Current studies on bilingualism have shown that all bilinguals—even proficient speakers—have activation during word recognition, where words from both languages are activated (e.g., Costa & Santesteban, 2004; Desmet & Duyck, 2007; Kroll & Stewart, 1994; Sunderman & Kroll, 2006). Unlike monolinguals, bilinguals use two languages in their everyday lives. At the same time, bilinguals show different patterns for negotiating competition across alternative responses, switching from one task to another, and ignoring irrelevant information, as observed from previous research (e.g., Bialystok, Craig, & Luk, 2012). Because there is interaction between the bilinguals' L1 and L2 on all representational levels (i.e., lexical, semantic, and syntactic), evidence suggests that it is impossible for bilingual speakers to process language in a purely monolingual mode (e.g., Desmet & Duyck, 2007). Similarly, when it comes to bilingual speech production, findings revealed that information from both languages are active, at least for a brief moment when bilinguals plan their speech (e.g., Costa, 2005; Kroll, Gullifer, McClain, Rossi & Martin, 2015).

One of the bilingual processing models that has been studied most extensively is the Bilingual Interactive Activation (BIA) model, an extension of McClelland and Rumelhart's (1981) Interactive Activation model for monolinguals (e.g., Kroll & Tokowicz, 2005; Van Heuven et al., 1998). The BIA model (e.g., Van Heuven et al., 1998) assumes that there is parallel activation and language non-selection between the lexical (word) and sublexical (phoneme) levels. According to the model, for languages that share similar orthographies, there will be parallel activation and competition at the lexical level, where the winner is activated and the others are suppressed (e.g.,

Van Heuven et al., 1998; Sunderman & Kroll, 2006). This includes cognates (words that share the same orthographical, phonological and semantic form), interlingual homographs (words that are similar in form but not in meaning), and orthographic neighbors (words that are only slightly different than the target word from each language) (e.g., Van Heuven et al., 1998; Kroll & Tokowicz, 2005). Therefore, if lexical access is non-selective, then a word that exists in both languages would be activated simultaneously and compete with each other until one form gets selected and the other suppressed. On the other hand, if lexical access is selective (where speakers can shut off one language completely), then the presence of the other form or meaning-related words from the other language should not have any effect on the target word. Instead, the bilingual speaker could process the chosen language like that of a monolingual speaker. In regard to the current language experiment, we want to investigate whether bilinguals' L1 influences their L2, especially during speech production, since the degree of activation and competition is strongly related to the likeness in words between the two languages (e.g., Van Heuven et al., 1998; Sunderman & Kroll, 2006; Green, 1998). Below are several key studies that support the view of parallel activation and language nonselective lexical access.

#### **Orthographic neighborhood effects in bilingual word recognition**

Van Heuven, Dijkstra, and Grainger (1998) found compelling evidence that supports the language nonselective lexical access. Van Heuven and colleagues conducted a series of experiments that investigated the effect of lexical neighbors on word recognition for Dutch English bilinguals. Lexical neighbors are words that differ from each other by just a single letter (i.e., *sofa* and *soda*). The first two experiments involved a word identification paradigm, also known as progressive demasking. The progressive demasking involves the target word being masked initially, before alternating with the mask, until the mask is removed all together.

In Experiment 1, participants (42 Dutch students) were divided into two groups, one for high English proficiency (HP) and one for low English proficiency (LP). The participants were presented with blocks of English or Dutch target words. Through a computer screen, participants were told to identify four-letter target words that would gradually appear on the screen. Followed by that, they were informed that there were two blocks, one with Dutch words (80 words) and the other with English words (80 words). The instructions were given either in Dutch or English. The target words appeared one at a time, through progressive demasking, and the participants must click on a button after they've identified the word before entering the word in a dialog box on the screen.

Results showed that target word recognition was significantly influenced by the number of orthographic neighbors in the non-target language. According to Van Heuven et al. (1998), there was inhibitory effect on words from both languages due to orthographic neighbors and cross language activation, which supported the hypothesis of non-selective activation. However, for HP bilinguals, there was more interference to English neighbors when the Dutch block was presented before the English block than the other way around, suggesting that high level bilinguals might be more influenced by their L1 than the low level bilinguals due to non-selective activation. Followed by the first experiment, another progressive demasking task was carried out in Experiment 2.

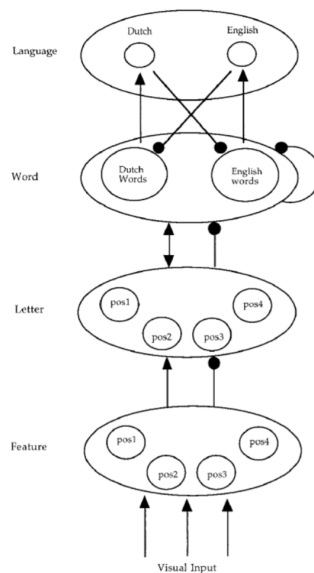
Experiment 2 further examined within- and between-language neighborhood effects in the progressive demasking task. The only difference between the first and second experiment is the use of mixed language presentation in the second one, with only one block of items in which both English and Dutch words were presented randomly. The results showed a significant inhibitory effect and reflected an increase of non-target language neighbors in both languages. At the same time, there appeared to be an increase in the effect of non-target language neighbors when going from a blocked to a mixed experimental design. This further confirmed the effect of non-target language neighbors influencing the identification of target words in the demasking task. In both experiments, only English target words had a facilitatory effect, while Dutch words showed significant inhibitory effect.

Experiment 3 examined whether the results obtained from the first two experiments with progressive demasking generalize to the lexical decision task. In addition, four nonword conditions with varying degrees of Dutch and English orthographic neighbors were also added to the experiment. Findings from this experiment showed that results from Experiment 1 and Experiment 2 did generalize to the lexical decision task. In order to further examine the role of language dominance, a fourth experiment was conducted.

Experiment 4 is a language-specific experiment and tested the English words used in the previous experiment with a new group of bilingual participants and an English monolingual control group. Both groups of participants were instructed to identify whether the string of four letter words were real English words. Results showed that while responses of English monolingual participants to English target words were not influenced by the number of Dutch neighbors, the

factors did significantly influence the response of the Dutch/English bilinguals. This further suggested that knowledge in the non-target language affects language processing.

The four experiments resulted in the language non-selective integrated-lexica model, also known as the Bilingual Interactive Activation (BIA) model (e.g., Van Heuven et al., 1998). The BIA model implements non-selective bottom-up processing and language-specific top-down processing. The model consists of a hierarchical arrangement of features, letters, words, and language nodes as shown in Figure 1 below.



**Figure 1.** An illustration of the Bilingual Interactive (BIA) model.

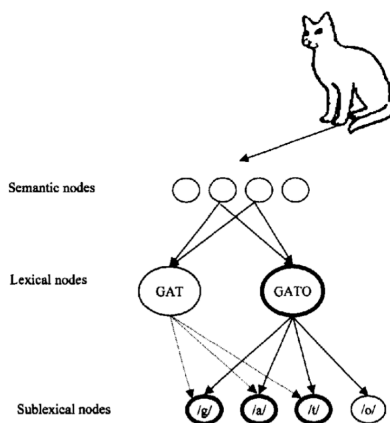
Based on the BIA model, when a proficient bilingual is presented visually with an input letter string, several lexical candidates are activated, regardless of the language. The activated lexical alternative then compete against each other until the winner is selected, while the others get suppressed (e.g., Van Heuven et al., 1998). At the same time, word recognition depends on the neighborhood size of the word in both languages. For instance, the speed with which Dutch-English bilinguals recognizes an English word like *farm* did not only depend on the number of English neighbors (e.g., *firm*, *fart*), but also on the number of neighbors in Dutch (e.g., *darm* [colon], *faam* [fame]) (e.g. Van Heuven et al., 1998; Dijkstra & Van Heuven, 2002; Desmet & Duyck, 2007). This further suggested that lexical orthographic representations from bilingual speakers' L1 is activated during word recognition in their L2. On the other hand, if there was language selectivity, or if a bilingual speaker could switch off one language completely, there



would be no effects for words with and without these cross-language form relatives, as observed in the above experiments. Nevertheless, the influence of L1 on L2 will be greater than the influence of L2 on L1, since most proficient bilingual speakers will maintain their dominance in one of their two languages (e.g., Van Heuven et al, 1998; Sunderman & Kroll, 2006).

### **Cognate facilitation and the cascaded activation model**

In another study on cognate facilitation, Costa, Caramazza, and Sebastian-Galles (2000) found substantial evidence that supports the claim of parallel activation, where both the selected and non-selected lexical nodes send activation to their phonological segments. Cognates can be interpreted as translation equivalents that have similar orthography and phonology. Bilinguals are usually at an advantage at recognizing cognates than monolinguals, since both their L1 and L2 are activated during word recognition. One example of cognate is the Catalan-Spanish word *cat* (e.g., *gat*—Catalan, *gato*—Spanish). It is predicted that the selection of phonemes corresponding to the target Spanish word *gato* is achieved very quickly due to cross language activations, since there is an overlap between the activated segments of the nonselected Catalan word *gat* (/g/, /a/, /t/) and the activated segments of the selected word *gato* (/g/, /a/, /t/, /o/) (e.g., Costa et al., 2000). Figure 2 below demonstrates the lexical and sublexical representation of Costa’s cascaded activation in processing cognate words.



**Figure 2.** Schematic representaton of lexical and sublexical access for cognate words.

Based on the cascaded activation model, two assumptions were made: (1) the bilinguals’ two languages share a common semantic system; and (2) there is parallel activation between the two languages. In order to test their predictions, Costa and colleagues (2000) carried out two experiments to investigate the performance of Catalan-Spanish bilinguals and Spanish monolinguals through picture naming tasks. For their experiments, the names were either cognates

or noncognates. If their predictions regarding cascaded activation is correct, there should only be difference between cognate and noncognate words for bilingual speakers.

Experiment 1 explored the extent to which picture naming is affected by cognate activation of the picture names. The participants were separated into two groups: highly proficient Catalan-Spanish bilinguals and monolingual Spanish speakers. Both groups were asked to name a set of picture in Spanish, where one set contained cognate names (e.g., *gato-gat* [cat]), while the other set did not (e.g., *mesa-taula* [table]). There were 40 pictures in total from a variety of semantic categories, and the frequency of the picture names had been manipulated to assess the power of the experiment and the comparability of the two groups of participants. The pictures were presented four times in four separate blocks. Participants were tested individually and had to name the pictures as quickly and as accurately as possible in Spanish. The experiment was self-paced and self-administered. There are three types of errors from the responses: (1) production of names that differed from the designated one for the picture; (2) verbal disfluencies; (3) recording failure (not from responses); and (4) words in which the participants named in Catalan.

Results revealed that bilingual and monolingual participants showed very different patterns for cognate status of pictures (cognate vs. noncognate), with bilingual speakers naming cognates faster than noncognates. Since this group of bilingual speakers were Catalan-dominant and responded in the target language Spanish, Costa et al. (2000) further investigated whether the magnitude of the cognate facilitation effect is proportional to the level of activation in the nonselected lexical node. A second experiment was conducted to compare the performance of bilinguals naming in their dominant language versus their nondominant language.

Experiment 2 used the same bilingual participants from the first experiment, who reported dominance for Catalan. The other group of participants were also highly proficient Spanish-Catalan bilinguals, but with Spanish as their dominant language. There were 80 pictures in total for the experiment, 40 pictures were used in the previous one. The onset of the picture names from both experiments were controlled, and the number of syllables in the cognate and noncognate conditions were also controlled. The procedure for the second experiment is the same as the previous, the only difference was the number of pictures included in each block and the randomization restrictions inside each block. The same type of error responses were also applied here as in Experiment 1.

Results indicated that the magnitude of the cognate effect was greater for the Catalan dominant group than the Spanish dominant group. It confirmed the results obtained from the previous experiment and suggested that the cognate status of words has a strong facilitation effect on bilinguals' naming performance. Costa and colleagues (2000) also demonstrated that cognate facilitation is controlled by the language of response, being larger when bilinguals respond in their nondominant language. The reason that there is a stronger cognate facilitation when bilinguals respond in their nondominant language is because translation in the dominant language has a stronger connection between semantic representation and their corresponding lexical nodes than in the weaker language (e.g., Costa et al., 2000). Results from both experiments further support the claim of parallel activation, where both the selected and nonselected lexical items activate their phonological segments.

### **Phonological overlap between L1 and L2**

Aside from orthography, evidence (e.g., Gollan, Forster & Frost, 1997) also suggests that there is phonological overlap between L1 and L2. In a study conducted by Gollan, Forster and Frost (1997), they examined whether the cognate effect may be purely phonological or whether it is the joint effect of overlap in both orthography and phonology. According to Gollan et al. (1997), one way to examine this issue is by removing orthographic overlap by using two languages that bear no similarity in script. For this reason, Hebrew was selected, since their script bears no visual resemblance to English. In order to examine the cognate effects between Hebrew and English, only cognates that overlap in phonological form were taken into account. If Hebrew-English bilinguals show stronger priming for cognates than noncognates, then it can be inferred that orthography plays no role in producing the cognate effect. The Hebrew cognates were considered "loan words" because many cognates were borrowed from English.

Experiment 1 tested Hebrew-dominant bilinguals with Hebrew primes and English targets (H-E). There were 40 Hebrew-dominant, Hebrew-English bilinguals, who completed a language-history questionnaire prior to the experiment. There were three lists of items, each consisting of 64 words and 64 nonwords. The nonwords were constructed by changing two letters of words matched in length to the targets in that list. For example, the Hebrew word for *obsession* is *obsesya*, and the nonword for that is *ogserya* (the stimuli were presented in Hebrew characters during the actual experiment). The items were presented through a computer screen, and the participants were instructed to press with their right hand if they saw a word, and press with their left hand if they

saw a nonword. The results showed that more priming was observed for cognates than for noncognates, but only for a limited subset of bilinguals. For some participants, it actually took longer to process the cognate targets in English (L2). According to Gollan et al. (1997), one explanation for the differences in cognate facilitation is due to the different levels of L2 within the participants. Through post hoc analyses, evidence showed that stronger priming effects for cognates were found in less balanced bilinguals.

Experiment 2 tested English-dominant bilinguals with the same three lists of items. Instead of having Hebrew primes and English targets, it was reversed with English primes and Hebrew targets (E-H). Experiment 2 followed the same procedure as Experiment 1. Results showed that there were clear priming effects in the within-language conditions and clear translation priming effects for both cognates and noncognates. However, there was much stronger translation priming effect for cognates than noncognates. One explanation could be because the bilinguals in Experiment 1 were more balanced than the bilinguals from Experiment 2. Results from both experiments indicated that superior priming for cognates can still be obtained even if the two languages have different scripts, since cognates are assumed to be jointly represented in both lexicons.

Experiment 3 and 4 aimed to further examine priming for cognates, and tested the priming effect from L2 to L1. In these two experiments, the primes were in L2 and the targets were in L1. The materials and procedures for these two experiments were the same as Experiment 1. Unlike the previous two experiments, results for Experiment 3 and Experiment 4 showed no enhanced priming for cognates, and no priming at all for noncognates. This indicated that bidirectional enhanced priming effects for cognates depend on the languages having similar orthographies (e.g., Gollan et al., 1997). Furthermore, results revealed that strong translation priming was only obtained with an L1 prime, not the other way around.

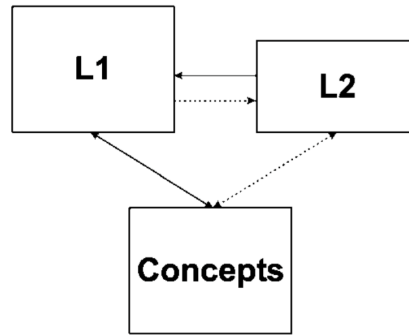
Gollan et al. (1997) found that enhanced priming for cognates was characteristic of processing for less proficient bilinguals through their study. This could be related to their heavy reliance on phonological code. Because accessing L1 targets does not rely on phonological recoding to the same extent as L2, this could explain why the priming effect was not bidirectional. At the same time, orthographic properties play a significant role in cross-language masked priming given that the results showed that Hebrew-English cognates were not accessed the same way as

cognates that share similar scripts (e.g., Gollan et al., 1997). Instead, enhanced cognate priming was only achieved by the shared phonological structure and emerged only when greater reliance on a phonological code is needed, as in the first two experiments from L1 to L2.

Based on the studies above, the BIA model assumes that word recognition is essentially non-selective with respect to language, given that word activation is affected by competing representations of both languages. The language nodes exert a top-down inhibitory effect on the words of the other language (as shown in Figure 2), and ultimately the winner language gets selected and the others suppressed. At the same time, even though speech planning starts at the conceptual level, evidence suggests that cross-language activation spreads and can reach all the way to the phonological level, and even beyond (e.g., Kroll, Gullifer, McClain, Rossi, & Martin, 2015). This is crucial to our current study on bilingual processing because when cross-language activation occurs, it could even affect the execution of speech. While the BIA model makes a number of claims regarding proficient bilinguals, it does not make specific claims about the effect of cross-language activation on L2 learners as they become more proficient in their L2 (e.g. Sunderman & Kroll, 2006). Another model of bilingual processing that will be discussed in contrast to the BIA model is the Revised Hierarchical Model (RHM), where it makes claims regarding language proficiency in bilingual processing.

### **The Revised Hierarchical Model (RHM)**

Another line of research that focused on the interlingual connections between lexical and conceptual representations of bilingual speakers is the Revised Hierarchical Model (RHM) (e.g., Kroll & Stewart, 1994; Sunderman & Kroll, 2006 ). The model focuses on how word-to-concept mappings are developed and accessed during language processing. The model also proposes independent lexical representation for words in each language, but an integrated conceptual system (e.g., Sunderman & Kroll, 2006). For example, during early stages of second language acquisition (SLA), the L2 words are hypothesized to be associated with their L1 translation, since only words in their L1 have direct access to meaning. Only after the bilingual speaker becomes more proficient in their L2, would their dependency to L1 diminish. Figure 3 below is a visual representation of the RHM.



**Figure 3.** The revised hierarchical model. (Adapted from Kroll & Stewart, 1994).

The evidence for RHM comes from experiments on translation performance (e.g. Kroll & Stewart, 1994; Talamas et al, 1999). According to the model, translation from L1 to L2 should be conceptually mediated, while translation from L2 to L1 should be lexically mediated. In Kroll and Stewart’s (1994) experiments, they tested proficient Dutch-English bilinguals’ ability to translate from both directions. Experiment 1 tested the category interference effect in picture naming (e.g., Kroll & Curley, 1988) when participants only responded with their first language (English). Subjects were briefly presented with pictures or words one at a time, and the items were blocked into lists that were either semantically categorized or randomly mixed. They were asked to name the words or pictures as quickly and as accurately as possible. Results showed that pictures produced category interference when they were named in a semantically organized list. On the other hand, words did not show sensitivity to the semantic context of the list, since evidence suggests that word naming reflects activity primarily at a lexical level of processing.

In Experiment 2, the participants were presented with the same lists of pictures or words as Experiment 1, but this time each list alternated between words and pictures from trial to trial. Subjects were instructed to name whichever stimulus appeared first and to expect an alternation from each trial. Results showed that word naming was much faster than picture naming, as shown in previous studies. However, the category interference effect diminished under the alternation conditions. This suggested that continuous access to related concepts produces increased activation at the concept level, which then causes interference due to difficulty in selecting one single lexical entry that best names the picture.

In Experiment 3, Kroll & Stewart (1994) wanted to determine whether category interference would occur in bilingual translation for highly proficient Dutch-English bilinguals. In order to do so, a Revised Hierarchical Model (RHM) was proposed. According to the model, both

lexical and conceptual links are active in bilingual memory, but the strengths of links differ due to translation asymmetry, with longer latencies in the L1 to L2 direction than in the L2 to L1 direction (e.g. Kroll & Stewart 1994; Kroll & Sunderman, 2006). For the naming task, subjects were asked to say the word that appeared on the screen aloud (English or Dutch), and for the translation task, they were instructed to translate the word on the screen to the other language (English to Dutch, or Dutch to English). Results showed that naming times were shorter than translation times, and it took longer to translate from L1 to L2 than the other way around. The findings further confirmed translation asymmetry, even for highly proficient bilinguals.

To summarize, while the BIA model predicts that nonselective access only exists for proficient bilinguals, it makes no predictions regarding less proficient bilinguals, where lexical activation occurs at the level of the word form. On the other hand, the RHM predicts that lexical activation occurs at the level of translation equivalents, and there is a shift on the reliance of L1 as the bilinguals become more proficient in their L2. Evidence further suggests that the degree of language activation depends on the types of lexical competitors at different levels of L2. In the study below, Sunderman and Kroll (2006) directly contrast the influence of lexical form and translation competitors in L1 to test the predictions of the BIA model and RHM in bilingual processing.

### **L1 activation during L2 lexical processing**

In Sunderman and Kroll's (2006) study, they wanted to investigate whether both types of form-related interference occur in different levels of bilinguals: (1) form related to the L1 word itself (as hypothesized by the BIA model); and (2) form related to the translation equivalent (as hypothesized by the RHM). The goal of their study was to assess to which degree do the two distinct types of lexical competitors activate during different levels of L2 proficiency, and to which extent does the conceptual information activate during lexical access at different levels of proficiency. While the RHM predicts that with increasing proficiency, L2 learners will be able to gain direct access to concepts in the L2 instead of relying on lexical links (translation equivalents), it is unclear at what stage this actually occurs. This is an extension of Talamas, Kroll, and Dufour's (1999) study, and the goal was to contrast and test the predictions made by the two models. Four questions were examined by Sunderman and Kroll (2006):

- (1) Is lexical information in the L1 active when L2 words are processed?

- (2) Does the activation of lexical information in the L1 differ for learners who vary in their L2 proficiency?
- (3) Does access to the meaning of L2 words increase with proficiency?
- (4) Does grammatical class function as a cue to lexical status and, if so, does it have the same effect on different levels of L2 learners?

Sunderman and Kroll (2006) recruited 107 university students with English as their L1 and Spanish as their L2. The students varied in their L2 proficiency and were divided into two proficiency groups—the less proficient group consisted of students who took three to six semesters of Spanish, while the more proficient group took seven or more semesters of Spanish. A language history questionnaire was distributed where the participants rated their own L1 and L2 proficiency in areas such as reading, writing, speaking and oral comprehension skills. The length of classroom exposure along with the self-rated results determined which group each participant was placed into.

In order to examine the participants' cognitive abilities, namely to assess whether or not there is a difference in cognitive abilities between low level and high level L2 speakers, several standardized measurements were used. The first behavioral measure was the reading span task adapted from Waters & Caplan (1996), which included both a storage component and a processing component. Eighty series of sentences taken from Waters and Caplan were presented to the participants through a computer screen. The participants had to judge the semantic plausibility of the sentences, and were instructed to make the judgment as quickly and accurately as possible. There were 80 sentences grouped into 20 sets with half being plausible and half implausible. The words from the sentences were controlled on several aspects, including word length, frequency, familiarity, concreteness and imageability. After a set of sentences, participants had to recall the final word of each sentence in that set and record it in a booklet. Accuracy was based upon how many accurate words the participants were able to recall from plausible sentences. The results showed that the difference between the two groups was not significant. Followed by that was the picture-naming task, where participants were presented with 40 dictionary-like line drawings and were asked to name the picture in Spanish as quickly as possible. Participants were instructed to name the picture aloud and their responses were tape-recorded. Results showed that the two groups differed significantly, with the more proficient group scoring a lot higher (58.6%) than the less proficient group (25.6%).



The actual experiment consisted of 48 correct translation pairs, one word in Spanish and the other in English. For each word pair, such as *cara-face*, there were a total of six distractors, two different distractors for each of the three conditions: form related to their first item, form related to the second item, and meaning related. In order to investigate sensitivity to grammar class, half of the distractors belonged in the same grammatical class as the target words, while the other did not. Participants were tested individually and were asked if two words were translation equivalents of each other. The word in their L2 always appeared first, followed by a word in English. Accuracy and speed were important for this experiment, and the speed of response was recorded to the nearest millisecond from the onset of the second word.

In terms of form-related lexical neighbors (e.g. *cara-card*), findings from this experiment showed that both groups of L2 speakers showed interference when the grammatical class was matched. Correlating to the predictions made by the BIA model, not only were the form-related relatives active in both languages during word recognition for proficient bilinguals, it was also active for the less proficient bilinguals. By comparing the activation in both level of proficiencies, results showed that the higher the proficiency, the quicker the response rate for the speakers.

In terms of form-related neighbors (e.g. *cara-fact*), the less proficient speakers were more sensitive to the form-related translation, while the more proficient speakers showed virtually no interference to this type of distracter. This finding correlated with the initial hypothesis proposed by the RHM, since the more advanced L2 speakers did not depend on their L1 to access the concept from their L2 words. However, despite of language proficiency, both groups showed interference when the distractors were not matched on grammatical class. Therefore, the results demonstrated that low proficiency bilinguals were sensitive to both form-related translation and grammatical class, while high proficiency bilinguals showed no interference and were faster at rejecting word pairs that were not equivalents (e.g., Sunderman & Kroll, 2006).

Lastly, results for meaning related pairs (e.g. *cara-head/cara-pretty*) suggested that both groups of speakers were affected by it. Furthermore, the meaning-related distractors showed that all participants were slower at rejecting meaning-related pairs, which suggested that both group of speakers were accessing conceptual information. Nonetheless, less proficient speakers were slower than the more proficient speakers at distractors, which was the same for all three conditions.

Through their study, Sunderman and Kroll (2006) confirmed the role of L1 in L2 processing in all levels of proficiency, and further suggested that bilingual speakers—despite of their proficiency—cannot shut off one language completely while speaking in another. On the other hand, L2 processing through L1 translation only occurred for the early/less proficient bilinguals, since the need for L1 translation diminishes as the speakers become more proficient in their L2. While the less proficient group were more sensitive to both lexical relatives and translation equivalents, both groups showed inhibitory to meaning-related distractors. This finding was significant because it showed that even less proficient speakers were sensitive to conceptual information during L2 processing, which contradicted the predictions made by the RHM (e.g., Sunderman & Kroll, 2006). At the same time, both groups showed sensitivity to grammatical class, which indicated that grammatical class might be able to influence processing even in an out-of-context task such as translation. While the results were in support of the claim that L1 is active during L2 processing for both low level and high level proficiency speakers, mediation of L2 through L1 only occurred during initial stages of L2 learning. Given that the BIA model and the RHM have very distinctive characteristics, neither model fully captured the complexity of lexical competition, especially when it comes to lower level phonological and phonetic proficiency. Aside from the BIA and RHM, another model concerning with bilingual speech production—the Inhibitory Control (IC) model—was investigated by Costa and Santesteban (2004) in the study below.

### **The Inhibitory Control Model**

#### **Lexical access in bilingual speech production**

Similar to comprehension, bilingual speech production also involves the activation of both the selected and nonselected alternatives in both languages (e.g., Costa et al., 2000; Costa & Santesteban, 2004; Hoshino & Kroll, 2008). There are two opposing views when it comes to lexical selection in bilingual speech production: the language specific (e.g. Dell, 1986) and the language nonspecific (e.g., Green, 1986). In the language specific view, the lexical selection mechanism is sensitive to specific properties of lexical nodes, and use them as guide for selection. For instance, in Dell's model (1986), the lexical selection mechanism is sensitive to grammatical class, where the selected item must match in grammatical class to the word the speaker wants to produce. On the other hand, the language nonspecific view assumes that the lexical selection mechanism is insensitive to the language intended to be expressed by the speaker. In this case, all lexical nodes

are activated until the right one gets selected. Previous studies (e.g., Costa et al., 2000; Costa & Santesteban, 2004) on picture naming demonstrated that naming latencies were faster with cognates than with noncognate names, given that both the selected and nonselected items activate their phonological segments. Gollan and Acenas (2000) argued that the cognate effect occurs because only the translation of the cognate words can send activation to the phonological elements of the target word, not the noncognate words. Therefore, the availability of phonemes would be higher for cognates due to the activation of both the selected and nonselected items, which supports the language nonspecific activation view.

The question under investigation in Costa and Santesteban's (2004) study concerns with whether or not the lexical nodes of the non-response language act as competitors. In order to examine whether inhibitory control accounts for lexical access in highly proficient bilingual speakers, Costa and Santesteban (2004) replicated the asymmetrical switching cost experiment by Meuter and Allport (1999). They also assessed whether L2 proficiency affects the pattern of language switching performance. Evidence from Meuter and Allport's study (1999) suggests that there is asymmetry in switching costs during bilingual speech, with greater inhibition for the dominant language (L1) when speaking in the less dominant language (L2). A post hoc analyses revealed that there was a direct correlation between the degree of asymmetry switch and L2 proficiency: the higher the L2 level, the smaller the asymmetry. These findings were initially in alignment with Green's Inhibitory Control (IC) model (1998). The IC model makes two critical assumptions: (1) inhibition is reactive and proportional to the level of activation of the words that are going to be suppressed; and (2) the more inhibition applied to a given lexicon (in this case L1), the harder it is to overcome such suppression on a subsequent trial (e.g. Green, 1998; Costa & Santesteban, 2004). More importantly, the IC model makes the claim that the amount of inhibition applied to one language is directly correlated to the bilinguals' proficiency in the other language. In order to assess how L2 proficiency level affects the processes of lexical selection in speech production, Costa and Santesteban (2004) compared language switching in highly proficient bilinguals and L2 learners through five picture naming experiments. In addition, they also examined mechanisms that control lexical access in bilingual speech production.

Experiment 1 replicated the asymmetrical switching costs (e.g., Meuter & Allport, 1999) in late L2 learners. Bilinguals from both groups were asked to perform language switching tasks

between their L1 and L2. Two groups of bilingual speakers were tested: (Group 1) Spanish-Catalan bilinguals with 1.5 years of learning experience in Catalan; and (Group 2) Korean-Spanish bilinguals, with 4 years learning experience in Spanish. There were 24 participants in total. For Group 1, 10 pictures of common objects with non-cognate names were used, and Group 2 used 8 of the 10 pictures with two new ones. Participants were required to select the language of response depending on the color of the picture (red or blue). The distribution of the color cue was assigned in a way that half of the participants were instructed to reply in their L1 for one color and L2 for the other, and the reverse was instructed for the other half. There were two trials: one with response language that matched the previous trial, and one with response language that was different from the previous trial. Each participant was tested individually, and had to name the pictures as quickly and accurately as possible. Results showed that “language of response” and “types of trial” were significant factors. The interaction between the two variables were also significant, and revealed that the degree of switching cost was larger for L1 than for L2 (e.g., Costa & Santesteban, 2004). This confirms the findings of Meuter and Allport (1999), where there was an asymmetrical switching cost depending on the response language. However, the response rate for non-switch trials was faster for L2 than L1 in Group 1, but the opposite was true for Group 2. In order to further assess the relationship between asymmetrical switching cost and the degree of language proficiency in L2 learners, a second experiment was conducted.

Experiment 2 examined language switching costs in highly proficient bilinguals. If the degree of the asymmetrical switching cost is correlated with L2 proficiency, then the reduction in the difference between L1 and L2 proficiency levels should also lead to a reduction in asymmetrical switching cost. Twelve native Spanish speakers with high proficiency in Catalan were recruited. The same materials and procedure were followed as the previous experiment. Results showed that highly proficient bilinguals experienced the same switching cost in L1 and L2, which differed greatly with the results from the first experiment (it was harder to switch to L1 than L2). A joint analysis revealed that the degree of asymmetrical switching cost was related to L2 proficiency, where highly proficient L2 speakers experienced less inhibition for L1. However, a slower rate for naming pictures in L1 was not expected from highly proficient speakers.

Experiment 3 further investigated the switching performance of highly proficient bilinguals. For this experiment, a larger set of pictures (40 instead of 10) was used to reduce the number of

receptions in each picture. Twelve highly proficient Spanish-Catalan speakers were recruited for this experiment, and none had participated in Experiment 2. There were forty pictures with non-cognates, and each picture only appeared once. Results from this experiment fully aligned with the ones from the previous experiment in two regards: (1) switching cost was the same from L1 to L2 as L2 to L1; and (2) naming responses was longer for L1 than L2. The first observation matches the pattern of the IC model, but the second observation suggested that highly proficient bilinguals developed a different type of selection mechanism. For instance, inhibition is not required in the non-response language when selecting words in the intended language for highly proficient bilinguals. In this case, the switching cost should be the same from L1 to L2 as from L2 to L1, which was observed through the result.

Experiment 4 tested whether highly proficient bilinguals show asymmetrical switching costs in their L1 and their weaker L3. If inhibitory control does explain the asymmetrical switching, then results should match that of Experiment 1. Twelve highly proficient Spanish-Catalan bilinguals who were learning English (L3) participated in the experiment. The same materials and procedure used in Experiment 1 were used here. The only difference is that participants performed the task in their L1 and L3. Results from this experiment did not match the results from Experiment 1, since the proficient speakers performed the switching task similarly in both their L1 and (much weaker) L3, without showing any asymmetrical switching costs. At the same time, the naming latencies were slower in L1 than L3, which replicated the results from Experiment 2 and Experiment 3. One explanation for the L2 advantage of L1 is that participants were biased toward the lexicalization of the weaker language. In order to further assess the L2 advantage over L1, a fifth experiment was conducted.

Experiment 5 examined whether the difference in naming latencies was caused by a bias in the naming process, in which the non-dominant language was prioritized at the beginning of lexicalization. According to Costa and Santesteban (2004), the bias can be eliminated if the participants know in advance which language the picture has to be named in. The same participants from Experiment 2 participated in this experiment. They were randomly assigned into two groups, corresponding to the two stimulus onset asynchronies (SOAs). Unlike Experiment 2, the language cue precedes the picture, and there was a difference in display time for the picture after the language cue onset. Results did not support the prediction, since the bias was not eliminated when

participants were informed about the language to be used before being presented with the picture. Furthermore, the differences between L1 and L2 naming latencies was not affected by the time given to the participants after the onset of the language cue. Therefore, a bias to start lexicalization in L2 does not provide sufficient evidence for L2 advantage in the language-switching task (e.g., Costa & Santesteban, 2004).

In summary, these five experiments demonstrated that there is a striking difference in language-switching abilities for L2 learners and highly proficient L2 speakers, where the less proficient group showed a clear asymmetrical switching cost (e.g., more inhibition for L1 when speaking in L2 than vice versa). However, results obtained from the highly proficient group did not align with the results found by Meuter and Allport (1999). The fact that highly proficient bilinguals were faster at naming the less dominant language suggests that there are different types of mechanisms involved in lexical access, which did not fully reflect the IC model. (e.g., Costa & Santesteban, 2004). Furthermore, results revealed that by becoming more proficient in both L1 and L2 could even affect the processing of a much weaker third language (L3). This suggests that lexical access in bilingualism is not limited to just the two languages that the bilinguals are proficient at, but also extends to other linguistic contexts. Despite of the findings from the five experiments, Costa and Santesteban (2004) could not find a single unitary explanation for why their results did not support the reactive inhibitory control proposed by Meuter and Allport (1999). Furthermore, the measures of proficiency used in these experiments were quite general and none of them measured lower level phonological and phonetic proficiency, which play a big role in distinguishing accent production between native and non-native speakers. Given that one of the most notable differences between L1 and L2 speakers is the difference in articulation in the target language, Chakraborty, Domsch, and Gonzales (2011) investigated the articulatory behaviors of nonnative speakers in their study below. More importantly, Chakraborty and colleagues examined how proficiency in L2 production is related to L2 processing skills among bilinguals of different L1 backgrounds.

### **Effect of L2 production on language proficiency** **Assessing accent proficiency in bilingualism**

When it comes to the production of English, there are several factors that separate native and nonnative speech. For instance, segmental and prosodic errors produced by L2 learners often affect comprehension of native listeners (e.g., Koster & Koet, 1993; Anderson-Hsieh, Johnson, &

Koehler, 1992). Furthermore, studies have shown that speakers with different L1 backgrounds can succumb to different production errors, such as L2 speakers of Chinese, Polish, and Spanish backgrounds tend to produce more errors with word-final stops (e.g., Flege & Davidian, 1984). Studies have also investigated the correlation between L2 accent, intelligibility, and proficiency, and showed that speakers with a strong non-native accent can still be highly intelligible and comprehensible. This indicated that L2 speech production does not necessarily correlate with comprehensibility (e.g., Munro & Derwing, 1995; Derwing & Munro, 1997). Unlike grammar and reading comprehension, it takes a much longer time for learners of L2 to achieve a native-like proficiency when it comes to accent production. One technique for assessing the differences in L2 proficiency is nonword repetition (NWR), which was first used by Gathercole and Baddeley (1990) to assess the speech production of children with language disorders. They found that NWR tapped phonological storage in working memory, and could provide a reliable indicator of language ability (e.g., Gathercole and Baddeley, 1990). Therefore, Chakraborty, Domsch, and Gonzales (2011) used NWR to examine how adult nonnative speakers' L2 production differ from native speakers', with a focus on the production of consonants in English words. The goal was to examine how nonnative speakers' L2 speaking proficiency is related to their L2 processing skills, as indicated by their accuracy of consonant production in different positions of the words.

Chakraborty et al. (2011) examined the L2 proficiency of native speakers of Bengali with English as their L2, given that Bengali is a language background that has not been heavily explored. There are several differences between the consonant inventory of English and Bengali. According to Chakraborty et al. (2011), English has six plosives (i.e., /p b t d k g/) without any aspirated counterparts which are contrastive, while Bengali has 20 plosives that include both aspirated and unaspirated contrastive counterparts. At the same time, English has labiodental fricatives /f/ and /v/ and approximates /w/ and /j/, which do not exist in the Bengali consonant inventory. Furthermore, in English, /p, t, k/ are aspirated by native speakers when produced in the initial position, but unaspirated when produced in the syllable final position, something that could become problematic for late L2 learners (e.g., Chakraborty et al., 2011). Due to the variations in consonant production between English and Bengali, Chakraborty et al. (2011) predicted that nonnative speakers (of English) with less exposure and proficiency in their L2 might have more difficulty producing those sounds, since they do not exist in Bengali. However, it might be less

problematic for L2 speakers to produce consonants that exist in both Bengali and English inventories. As a result, the study also investigated whether a production difference between the high and low L2 proficiency groups is observable when common consonants are also included in the stimuli.

For this experiment, three different groups of subjects were recruited: 12 normal adult Bengali-English bilinguals with early exposure to Indian-English (who had received primary and secondary education in English), 12 normal adult Bengali-English bilinguals with late exposure to Indian English (who started learning English from college), and 12 normal monolingual adult English speakers. Before testing the participants' accent production, they first completed the Speaking Grammar subtest of TOAL-3, which reflected the participants' knowledge of English syntax. The maximum attainable raw score for the subtest is 30, and results showed little statistical significance between the early bilingual group and native English speaking group, while the late exposure group scored significantly lower (Chakraborty et al., 2011). In order to clearly identify the two groups of bilinguals, the participants with the top 10 scores were selected from the early-exposed group and the lowest 10 scores were selected from the late-exposed group. These two groups will then be compared with the top 10 scores from the native English speaking group. In this case, the age of exposure correlated with the test results.

Chakraborty et al. (2011) used nonword targets as stimuli to test the participants' production, which were originally developed by Dollaghan and Campbell (1998). There were 16 nonword stimuli with four different syllable lengths that were incorporated into 96 phonemes, and all the nonwords had consonants in word-initial and word-final position. A monolingual speaker of American English recorded the target stimuli and stored the data in a computer. For the nonwords test, the words came out in a fixed pattern, with the monosyllabic nonwords playing first, followed by two-, three-, and four-syllable words. After each nonword has been played, the participants were asked to repeat after what they heard, and each nonword only gets played once. The nonwords were transcribed by a certified speech-language pathologist. After all the recordings were transcribed, the nonwords were scored for accuracy. The consonants were compared to the target phonemes, and the scores were recorded as correct or incorrect. Any substitutions and omissions in the recording was considered as incorrect.



Results for the NWR showed that the “early/high” and the monolingual groups were comparable in their results, while the “late/low” group was statistically lower. According to the data, all three groups scored the lowest for monosyllabic words, with the monolingual group scoring the highest. Both late and early exposure groups scored similar for monosyllabic words, which most likely indicated that a one-syllable nonword might not differentiate the two levels of L2 proficiency when they share the same Bengali L1. However, the early exposure group preformed much better on the two-syllable nonwords, and all three groups had similar accuracy scores for the two-, three, and four-syllable targets. The scores showed that early-bilinguals were similar in their production and accuracy as native American English speakers, while the late-bilinguals scored significantly lower. The biggest difference occurred for the production of consonants in word-final position, where the late-bilinguals had a much lower accuracy compared to the other two groups (Chakraborty, et al., 2011).

The results from these experiments confirmed that the L2 experience (such as amount of exposure) influenced the overall accuracy of consonant production, and also confirmed that age of exposure to L2 influenced production accuracy. Chakraborty and colleagues (2011) also demonstrated that the production of nonwords reflected the language processing ability of the speakers, and also revealed that early exposure to a dialectal variation of English appeared to be phonologically advantageous for production of a different dialect of the same L2. For instance, there is a different dialectal difference between native English speakers and proficient Bengali-English speakers, even though both showed the similar degree of comprehensibility during production. Therefore, the difference between the two L2 groups was in their L2 experience, which suggested that early exposure to the L2 is advantageous in obtaining a more native-like accent (Chakraborty, et al., 2011).

These results reveal crucial information regarding the relationship between language experience and language proficiency (i.e., length of exposure, age of acquisition, how L2 was acquired, etc.), which are important language variables that will be investigated under our current study. While evidence shows that the ability to accurately produce nonwords reflected the bilinguals’ processing ability, it does not take the role of verbal fluency into consideration, which involves the production of complete sentences. In a more recent study conducted by De Jong, Groenhout, Schoonen, and Hulstijn (2013), they examined whether the ability to produce fluent

speech is a reflection of the bilinguals' L2 proficiency, or merely a reflection of the speakers' speech style.

### **Is L2 verbal fluency an accurate indicator of language proficiency?**

According to De Jong and colleagues (2013), the term fluency in the simplest form describes the speed and smoothness of speech production, without pauses, repetitions, and repairs. When it comes to L2 testing, speaking fluency is used as a variable in the evaluation of speaking proficiency, and it is usually judged by the temporal aspects of speech (e.g., Shriberg, 1994; De Jong et al, 2013). However, the question at hand is whether the L2 speakers' speech fluency is really a reflection of their language ability, or simply a reflection of their speaking style (such as having frequent pauses even when speaking in their L1). A recent study conducted by Derwing, Munro, Thomson, and Rossiter (2009) suggests that there is a significant correlation between L1 and L2 behavior in terms of the number of pauses per second, speech rate, and pruned syllables per second for Mandarin and Slavic speakers of English. This indicated that speakers who are not fluent by default could carry their L1 speech characteristics to their L2 speech production, which makes L2 proficiency judgements (based on verbal fluency) difficult to validate (e.g., De Jong et al., 2013). One way to avoid misjudging L2 verbal fluency is by gathering both L1 data and L2 data, in order to take bilinguals' L1 fluency into account (e.g., Segalowitz, 2010). By doing so, disfluencies that are specifically related to the use of speakers' L2 can be separated from disfluencies that appear in their L1.

In their study, De Jong and colleagues (2013) examined L2 *utterance fluency* in bilingual speakers. Utterance fluency is defined by the number of filled pauses, corrections, and repairs that occur during L2 speech production, and by measuring the duration of pauses. Utterance fluency will be examined in two ways: uncorrected measures and corrected measures that take L1 behavior into account. Both measures will be tested against an objective measure of L2 proficiency (vocabulary knowledge), in order to examine whether the corrected measures can better reflect L2 *cognitive fluency* (the ability for L2 speakers to smoothly translate thoughts into speech). For this study, two different groups of L1 participants (English and Turkish) were recruited to test against their common L2 (Dutch). De Jong et al. (2013) chose bilingual speakers of English and Turkish background for a more distributed result, given that English is typologically close to Dutch, and Turkish is typologically distant from Dutch. Three key issues were investigated in this study:

1. To what extent can different measures of L2 fluency be predicted from the equivalent measures in L1?
2. Are L2 fluency measures that are corrected for L1 fluency better predictors of L2 proficiency?
3. Is the predicative value of corrected measures of L2 fluency dependent on typology between L1 and L2?

The participants consisted of 29 native English speakers and 24 native Turkish speakers. Both groups of participants moved to the Netherlands between the ages of 18 and 40, and maintained an intermediate to advanced level of their L2, Dutch. In order to assess the overall L2 proficiency separately from L2 speaking proficiency, a productive vocabulary task was given for the objective language measurement (e.g., De Jong et al., 2013). In the paper-and-pencil-task, knowledge of single words and multiword units were elicited. There were 90 items for single words and 26 items for multiword units. For each item, a sentence was presented with the target word omitted except for the first letter(s). However, more letters were presented when alternative words beginning with the same letter could also be used. In order to maximize testing productive knowledge of the targeted word, De Jong et al. (2013) purposely selected words with higher frequency in the carrier sentence than that of the omitted word. Aside from assessing participants' productive knowledge, this task also examined their comprehension skills. Unlike single words, the multiword units consisted of prepositional phrases and verb-noun collocations. For each unit, the preposition or main verb was omitted and the participants must fill in the gap. No first letter(s) were given for the multiword units.

In order to elicit L2 speech proficiency, the language experiment used the eight speaking tasks in De Jong, Steinel, Florijn, Schoonen, and Hulstijn's (2012) study. All of the speaking tasks matched the two languages in three aspects: complexity, formality, and discourse mode. For example, all of the tasks from Table 1 performed by the speakers in their L2 (Dutch), a mirroring set of questions were presented in English for the L1 English speakers. The Turkish tasks were translated from the English tasks for the L1 Turkish speakers. Native-speaking experimenters were present during each set of tasks (i.e., a native Dutch-speaking experimenter was present during their L2 tasks). After the completion of the vocabulary task, the participants performed their L2 (Dutch) speaking tasks, followed by their L1 (English or Turkish) speaking tasks 1 to 4 weeks

later. All three tasks took place during different sessions. The speech recordings were transcribed and annotated by two research assistants, and information relevant to fluency measures were also reported.

**Table 1.** Speaking tasks adapted from De Jong et al. (2012) that was performed in Dutch.

<b>Task 1</b> (simple, informal, descriptive)	Participants speaks on the phone to a friend, describing the apartment of friend who have recently moved house
<b>Task 2</b> (simple, formal, descriptive)	Participant, who has witnessed a road accident some time ago, is in a courtroom, describing to the judge what had happened.
<b>Task 3</b> (simple, informal, argumentative)	Participant advises his/her sister on how to choose between (or combine) child care, further education, and paid work.
<b>Task 4</b> (simple, formal, argumentative)	Participant is present at a neighborhood meeting in which an official has just proposed to build a school playground, separated by a road from the school building. Participant gets up to speak, takes the floor, and argues against the planned location of the playground.
<b>Task 5</b> (complex, informal, descriptive)	Participant tells a friend about the development of unemployment among women and men over the last ten years.
<b>Task 6</b> (complex, informal, argumentative)	Participant discusses the pros and cons of three means of transportation (public transportation, bicycle, automobile) on how to solve the problem of traffic congestions.
<b>Task 7</b> (complex, formal, descriptive)	Participant works at the employment office of a hospital and tells a candidate for a nurse position what the main tasks in the vacant position are.
<b>Task 8</b> (complex, formal, argumentative)	Participant, who is the manager of a supermarket, addresses a neighborhood meeting and argues which one of three alternative plans for building a car park he/she prefers.

According to De Jong et al. (2013), fluency measures were calculated for three different aspects of fluency: *speed fluency* (mean syllable duration in L1 and L2), *breakdown fluency* (mean length of silent pauses within and mean length of pauses between analysis of speech units (ASU)), and *repair fluency* (number of repetitions per second speaking time and number of corrections per second speaking time in L1 and L2). Furthermore, all frequency measures of fluency were divided from the total speaking time, excluding pausing time.

Results for this experiment showed that the L1 Turkish and L1 English speakers performed equally well in their L2 Dutch on the vocabulary task. The Shapiro-Wilk measures were used to test whether L1 and L2 measures could be assumed to be normally distributed (e.g., De Jong et al., 2013). In terms of speech fluency, De Jong and colleagues (2013) observed that both groups were less fluent in their L2 than their L1, with the largest difference for mean syllable duration and smallest difference for number of repetitions. At the same time, even though L1 Turkish and L1

English speakers showed variance in aspects of speech production in their own language (i.e., syllable duration, silent pauses per second, and repetitions), no difference was observed in those areas when speaking in their L2. In terms of predicting L2 fluency, the experimenters computed all fluency measures within the L1 and L2. Data suggested that even though there are large differences between the measures of L1 fluency between the two L1 languages (English and Turkish), the relationship between L1 and L2 is the same for both groups of speakers. At the same time, evidence indicated that all measures of L2 fluency can be predicted on the basis of the L1 fluency measurement alone, up to a certain extent. Lastly, data revealed that correcting the L2 scores of syllable duration for L1 behavior led to more accurate predictions of vocabulary scores, with no significant differences found in other fluency variables (e.g., De Jong et al., 2013). De Jong and colleagues (2013) concluded that utilizing corrected measures of L2 fluency (by measuring both L1 and L2 speech) will benefit research in L2 production, since judging L2 speakers' proficiency on the uncorrected measure alone is partly judging them on their personal speaking style, which is unrelated to their L2 proficiency.

Data from the study above suggested that bilinguals' L1 fluency does play a role in influencing their L2, especially when it comes to speech production. This is an important piece of evidence, given that our current study also tries to examine whether the bilinguals' L2 speech proficiency reflects aspects of their L1 speech, by examining their attitude and reference for their L1. Furthermore, when it comes to self-rated speech proficiency, bilingual speakers might not be aware of the errors that exist in their L2 speech, since some of those errors could be a reflection of their L1 speaking style. Since this study concerns with self-rated language proficiency, it is crucial to understand why self-assessments have proven to be an effective tool in assessing bilingual proficiency. The next section will focus on the role of self-assessed language proficiency through two studies, which served as the foundation of our current study (e.g., Marian et al., 2007; Delgado et al., 1999). However, no previous studies in bilingual language processing has explicitly investigated the role of self-rated accent proficiency and its relationship with other aspects of L2 proficiency, which makes the current study novel in its field of research.

### **Self-rated language profile**

#### **Self-Assessment of Linguistic Skills**

One of the main goals of the current language experiment is to examine whether the self-ratings of language proficiency from bilingual speakers is credible and can reflect their actual

language skills. A number of studies (e.g., Delgado et al., 1999; Marian et al., 2007) have shown that self-rated language proficiency is robust and does reflect the overall language proficiency of the bilingual speaker.

Delgado and colleagues (1999) examined the role of self-assessment in their study on self-rated linguistic skills by Spanish English bilinguals. In their study, bilinguals rated their skills in English and Spanish and were given an objective test to measure against their answers to their language proficiency. It was predicted that participants would be fairly accurate in accessing their skills in both languages. At the same time, it was predicted that self-assessments would be more accurate after feedback than prior to feedback, and ratings for language proficiency would vary depending on the language the questions were presented in.

Eighty Spanish-English bilinguals were recruited for the experiment. A picture-naming task consisting of 24 line drawings was distributed to the participants prior to the experiment, since it was found to distinguish skilled from unskilled language users. Afterwards, a biographical questionnaire regarding their ethnicity, education, and language background was given. Finally, participants rated their language skills on a scale from one to five, one being very poor ability, and five being excellent ability. These ratings were based on how they behave during everyday activities. For the objective measure, the Spanish and English versions of the Woodcock-Muñoz Language Survey was used to test their bilingual abilities. Both versions contained four subtests: Picture Vocabulary, Verbal Analogies, Word Identification, and Dictation. The mean of all the subtests provided a general measure of each language. After the initial picture naming task, the qualified participants were assigned randomly to one of four Hispanic examiners for individual testing. All instructions and materials were given in English. After the self-assessed section, they were given the Woodcock-Muñoz survey in both English and Spanish. Finally, the participants were asked to rate for a second time on all 10 self-rating scales after receiving feedbacks on their performance. The rating scales were written in Spanish for half of the participants (by two Spanish speaking examiners), and English for the other half (by two English speaking examiners).

In terms of self-ratings, results showed that the bilinguals initially overestimated their overall fluency for both languages, but after receiving indirect feedbacks, their post self-ratings were significantly lower. Delgado and colleagues (1999) found certain discrepancies between ratings for Spanish and ratings for English, since results revealed that the participants judged their

speaking and listening skills to be better in Spanish, while their reading and writing skills to be better in English. However, the interaction between language and presentation was not significant. When tested against objectively measured tasks, the results revealed that Spanish language skills were estimated more accurately than English skills. For English, only evaluations for reading-writing skills correlated with the Woodcock-Muñoz subtests, but all correlations between Spanish self-assessments and Spanish test scores were significant. Therefore, these results showed that participants were accurate in assessing their overall Spanish skills, but not their overall English skills (e.g., Delgado et al., 1999). At the same time, self-evaluations were better after taking the objective test in four cases for Spanish, but no significant improvement was found for English.

According to Delgado et al. (1999) one explanation for why the bilinguals were better at judging their Spanish skills than their English skills was because the participants learned Spanish first (according to the demographic information), and had early exposure to Spanish at home. Early exposure may have provided the bilingual speakers with sufficient practice and feedbacks regarding spoken Spanish, so they were able to rate their oral Spanish skills more accurately. On the other hand, the participants probably received more feedbacks on reading and writing than for speaking in English, while they were in school. Furthermore, questions should be more specific regarding the use of each language, such as indicating where and how the language was used, rather than generalizing to “everyday activities” (e.g., Delgado et al., 1999). The results demonstrated that Spanish English bilinguals were fairly accurate in evaluating their reading-writing skills in Spanish, but somewhat less accurate in assessing those skills in English. When it comes to oral judgment, only self-assessed Spanish oral skills proved to be fairly accurate, while self-rated English oral skills did not prove to be reliable. Evidence from previous studies (e.g., Bahrick et al., 1994) also suggested that the relationship between self-rated language skills and objective tasks was weakest for oral comprehension. Nevertheless, the self-assessed questionnaire was able to capture the bilinguals’ overall L1 proficiency, and part of their L2 proficiency, thus providing credibility to self-assessments.

### **LEAP-Q**

Another study that examined self-assessed language proficiency is the Language Experience and Proficiency Questionnaire (LEAP-Q), created by Marian and colleagues (2007) to assess language profiles in bilinguals and monolinguals. While evidence from previous studies demonstrated the robustness of self-rated language proficiency, it also yielded inconsistent

findings in a number of areas, such as lexical processing, and processing regarding phonology and orthography (e.g., Marian et al., 2007). The LEAP-Q was created to provide a more comprehensive, valid, and reliable questionnaire across bilingual populations, and capture factors that had been previously considered significant to the bilingual status. These contributors can be divided into three major categories: language proficiency, language dominance, and language preference. According to Marian and colleagues (2007), the LEAP-Q elicited proficiency ratings in speaking, listening, reading, and writing, rather than combining them with different performance domains into a cumulative domain. For instance, proficiency ratings from the LEAP-Q presented different predicative information for different linguistic skills, and participants indicated their dominance in a scale from 0-10, with 0 being the lowest and 10 being the highest. The questionnaire also included questions that examined the bilinguals' language dominance in both global (dominance) and specific (preference) aspects for both L1 and L2. Furthermore, the LEAP-Q elicited current language exposure from the bilinguals, such as interaction with family and friends, exposure to different types of media outlets, and independent learning among others. Table 2 below shows the categories and items under the language history questionnaire from the LEAP-Q in Study 1.

**Table 2.** Self-reported language history and proficiency adapted from LEAP-Q (2007)

<b>Self-reported proficiency</b>	<b>Age milestones (years)</b>	<b>Immersion duration (years)</b>	<b>Contribution to language learning</b>	<b>Extent of language exposure</b>	<b>Self-report of foreign accent</b>
Understanding Speaking Reading Writing	Started learning Attained fluency Started reading Attained reading fluency	In a country In a family In a school	From family From friends From reading From TV From radio From self-instruction	Family Friends Reading TV Radio Independent study	Perceived by self Perceived by others

The LEAP-Q was developed as a tool to examine the bilingual's experience and proficiency profiles in both L1 and L2. Two experiments were conducted to examine both the internal validity of the LEAP-Q (self-assessment) and the criterion-based (standardized proficiency measures) validity. Study 1 examined the internal validity of the questionnaire by analyzing the responses of 52 bilinguals from a diverse language background through factor analysis and multiple regression analysis. Through the factor analysis, eight factors were yielded from the responses and a name was assigned for each factor that reflected the underlying construct



of their components. Table 3 below shows the results of the factors yielded in study 1. The results from Study 1 led to the exclusion of four measures from the questionnaires: measures of comfort across modalities, writing proficiency, current classroom exposure, and percentage of bilingual contacts. These measures were eliminated because they either yielded similar values to other measures of proficiency, or did not correlate with any other measures and did not load onto any factor in the factor analysis.

**Table 3.** Factors yielded in Study 1.

Factor 1: L1 Competence	Loading values	Factor 2: Late L2 Learning	Loading values	Factor 3: L2 Competence	Loading values	Factor 4: L1 Maintenance	Loading values
Proficiency reading	.947	Age became fluent	.864	Exposure (% time)	.923	L1 exposure to classes	.916
Comfort understanding	.910	Age began acquiring	.859	Exposure to TV	.908	L1 exposure to TV	.914
Proficiency understanding	.910	Age became fluent reader	.855	Exposure to friends	.861	L1 exposure to radio	.831
Comfort writing	.903	Comfort understanding	-.803	Exposure to radio	.772	L1 exposure to reading	.776
Proficiency writing	.896	Age began reading	.751	Writing proficiency	.660	L2 learning from reading	-.727
Comfort reading	.884	Proficiency understanding	-.697	Exposure to family	.621	L1 exposure (% time)	.627
Identified accent	-.788	Years in a country	-.681	Comfort writing	.592	L1 exposure to friends	.530
Comfort speaking	.748	Learning from tapes	.601	Preference to speak	.590		
Proficiency speaking	.704	Proficiency speaking	-.580	Exposure to reading	.564		
Cultural identification	.526			Exposure to classes	.543		
Perceived accent	-.517			Learning from reading	-.519		
Preference to read	.457						
% variance	23.480		13.383		9.625		7.534
Cumulative variance	23.480		36.862		46.488		54.021
Factor 5: Late L2 Immersion	Loading values	Factor 6: Media-Based Learning	Loading values	Factor 7: Non-Native Status	Loading values	Factor 8: Balanced Immersion	Loading values
L1 years of class learning	.728	L1 learning from TV	.866	L2 perceived accent	.839	L1 learning from friends	-.813
L2 years in workplace	.725	L2 learning from TV	.838	L2 identified accent	.615	L2 years of schooling	.627
L1 years in workplace	.714	L1 learning from the radio	.741	L2 cultural identification	-.602	L1 years in a family	.622
Proficiency reading L2	-.687	L2 learning from the radio	.652	L1 age became fluent	.590	L2 years in a classroom	.541
L2 learning from friends	-.683	L2 comfort reading	-.476	L2 learning from family	-.519	L1 years in a country	.499
L2 learning in a classroom	-.556						
L1 years in school	.476						
% variance	6.424		6.226		5.049		4.232
Cumulative variance	60.445		66.671		71.720		75.952

For Study 2, Marian et al. (2007) used the revised LEAP-Q to confirm the internal validity of the LEAP-Q and to establish criterion-referenced validity by comparing the results of self-rated and standardized proficiency measures. In Study 2, 50 English and Spanish bilinguals were recruited. The participants first completed the revised version of the questionnaire on a computer for the first part of the experiment. Upon completion of the questionnaire, they were administered a battery of standardized tests to measure their language ability. Table 4 shows the seven behavioral measures that were tested:

**Table 4.** List of behavioral measurements and description adapted from Study 2.

Behavioral tasks	Response required
Reading Fluency	Participants are required to read as many sentences as possible within a 3-min interval and to decide whether each sentence was true or false
Passage comprehension	Participants are required to read passages and supply missing words

Productive picture vocabulary	Participants are required to name pictures
Oral comprehension	Participants are required to listen to passages and supply missing words
Sound awareness	Participants are required to complete a rhyming task, a sound deletion task, a sound substitution task, and a sound reversal task
Receptive vocabulary test	Participants are required to identify picture in response to auditory instructions
Grammaticality judgement test	Participants are required to read 50 sentences in English and 50 sentences in Spanish and to judge whether they are grammatically correct

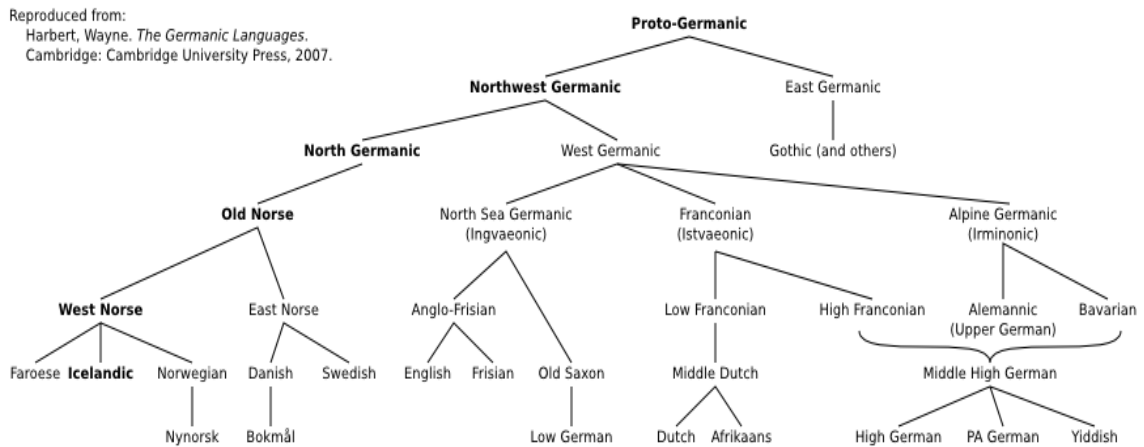
Results from Study 2 revealed positive correlations between the standardized tests and self-reported measures of proficiency on several different aspects. For instance, participants reported higher levels of L1 proficiency than their L2 proficiency in the questionnaire, especially in areas such as understanding and reading. This was reflected through the participants' behavioral measures through the battery of tasks, where they performed significantly lower in L2 than in L1. Furthermore, results suggested that measures of comprehension and grammatical judgements were the best predictors of self-reported proficiency for both L1 and L2 (e.g., Marian et al., 2007). While the participants rated themselves highly for L2 reading proficiency, they showed poor performance on behavioral tasks, which indicated that they had overestimated their true reading skills. In terms of factor analysis, results for Study 2 revealed that questions that were expected to measure the same underlying construct did cluster together. Furthermore, the multiple regression analysis suggested that self-reported measures of language proficiency were indicative of their linguistic performance to a certain degree. The findings in Study 2 ultimately suggested that language history profiles can predict standardized performances to a certain degree, but vary across tasks. Therefore, the LEAP-Q is an effective and reliable tool for assessing bilingual language status, and includes multiple aspects of the bilinguals' language proficiency.

While evidence from the LEAP-Q suggests strong correlation between self-assessed language proficiency and linguistic performance, it did not take language attitude and language preference (for both L1 and L2) into consideration. Even though the participants were asked to rate different areas of their L2 proficiency, it did not investigate the role of self-rated accent proficiency, such as how much effort they put into sounding native-like and how important it is for them to have a good L2 accent. Therefore, more items will be included in the current language profile, especially in regard to accent-related variables. In addition, more items will also be included for the behavioral measures, in order to provide a more comprehensive reflection of the

bilinguals' language proficiency (i.e., speaking, understanding, reading, writing, grammar, vocabulary, and spelling). Given that the current study investigates language processing among highly proficient Norwegian-English bilinguals, it is crucial to understand some of the key structural differences between the two languages. The next section will provide a brief explanation on the key differences in phonology, morphology, syntax, and prosodic features between English and Norwegian.

### Norwegian English Language Comparison

While English and Norwegian are both Germanic languages with a common Proto-Germanic ancestor, the two have evolved and separated into its own over time. The diagram below is a language tree that demonstrates the languages and language changes since Proto-Germanic.



**Figure 4** An illustration of the evolution of Germanic languages. Image extracted from Cambridge University Press, 2007.

Given that the degree of activation and competition depends on how much influence each language has on the other (e.g., Costa & Santesteban, 2004; Sunderman & Kroll, 2006; Van Heuven et al., 1998), this section will explore some of the distinctive features that set the two languages apart. Some of the topics that will be discussed in this section include the differences in the phoneme inventory (vowel and consonant), differences in syntax and morphology, as well as differences in tone and stress between the two languages. These distinctions form the basis of our standardized measurements, where we aim to capture the most comprehensive reflection of the bilinguals' linguistic ability.

### Cognates and homographs

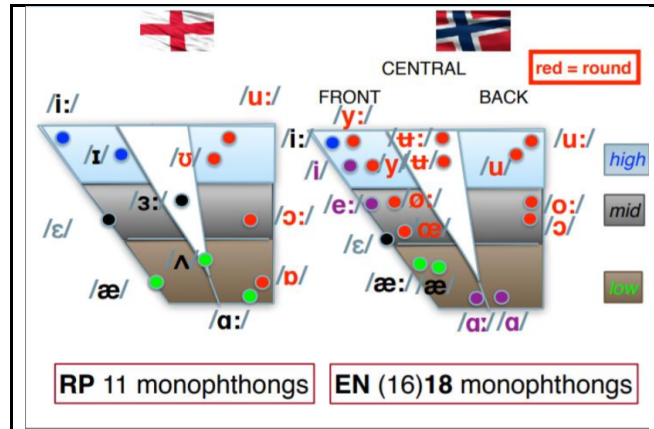
As discussed earlier, in languages that share the same alphabet, it is possible to have words that are *cognates*, i.e. they have the same (or nearly identical) spelling, pronunciation, and meaning (e.g., Kroll, 2008). For instance, words such as “hotel,” “student,” and “toilet” are cognates between English and Norwegian, since they are spelled the same and carry similar meanings in both languages. On the other hand, there are also *homographs*, which are words that are spelled the same but have no relation in meaning (see Table 5). While cognates help facilitate reading (by making reading faster), homographs do the opposite (by slowing down the word recognition process), also known as inhibition (e.g., Poort & Rodd, 2017; Costa, Caramazza & Galles, 2000; Van Heuven et al., 1998).

**Table 5.** Examples of cognates and homographs in English and Norwegian

Norwegian Word	English Meaning	Norwegian Meaning
pen	a writing utensil	pretty
gift	a present	married
far	long distance	father

### Phonetic Inventory: Vowels

Even though English and Norwegian stemmed from the same Germanic branch of the Indo European family, the two languages evolved over time (see Figure 4 for reference), with one belonging to the West Germanic branches (i.e., English) and one belonging to the North Germanic branches (i.e., Norwegian). For this section, I will be referring RP (Received Pronunciation) for English (typically recognized as British English), and EN (Eastern Norwegian) for Norwegian (typically recognized as the Oslo dialect). In the English alphabet, there are six vowels that represent 11 monophthongs in RP. On the other hand, there are nine vowels in the Norwegian alphabet (with the addition of æ, ø, and å) that represent 18 monophthongs for EN. Figure 5 below is a comparison between the vowel inventory in RP and EN.



**Figure 5.** A comparison of English and Norwegian vowel inventory. (Wetterlin, 2017)

When comparing the English and Norwegian vowel inventory side by side, a number of differences can be observed. The most interesting difference is that Norwegian has front rounded vowels but English does not. Some other differences include the lack of “lax” vowels /ɪ, ʊ/ in EN (i.e., *bitt* vs. *bit*), as well as the lack of mid and low central vowels /ɜ:/ (i.e., *nurse*) and /ʌ/ (i.e., *duck*) (Wetterlin, 2017). Over the course of time, changes have occurred in the vowel inventory in both languages, thus distinguishing the two languages phonemically, and may cause difficulty for L2 learners to produce.

### **Phonetic Inventory: Consonants**

As in the case of vowels, there are also differences in the consonant inventories between the two languages. While RP and EN have the same number of consonants (24 consonants from each language), the inventories are different (see Figure 6). Compared to English, the Norwegian consonant inventory lacks voiced labiodental fricatives /v/, dental fricatives /θ, ð/ (voiceless and voiced), voiced alveolar fricative /z/, and postalveolar fricatives /ʃ, ʒ/ (both voiceless and voiced). In addition, there are no postalveolar affricates /tʃ, dʒ/, and approximant /r/ in Norwegian (also known as the “rolling-r”). Below is an illustration comparison of the English and Norwegian consonant inventory.

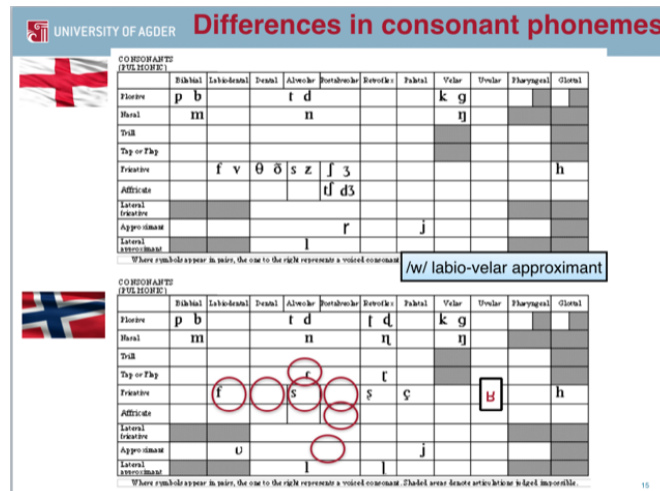


Figure 6. A comparison of English and Norwegian consonant inventory. (Wetterlin, 2017)

As a result, despite of having the same number of consonants in each inventory, the inventories themselves are not the same, and certain phonemes can become problematic for L2 learners (especially those with lower English proficiency) to pronounce. In our current study, we designed tasks that specifically target the sounds that are relatively foreign to native Norwegian speakers. From Figure 6 above, we see that English fricatives might be the most difficult for Norwegian speakers to produce or differentiate, especially for the voicing phonemes. Table 6 below is a chart used for teaching Norwegian students that highlights the difficulties in consonant production for L2 speakers of English.

Table 6. An illustration of English consonants that cause difficulty for L1 Norwegian speakers. (Wetterlin, 2017)

New sounds - for most Norwegian dialects	
/ð θ/	/ð/ there is the /θ/ three thistle theatre
/s z/	seven /s/ singing /z/ zebras
/ʃ ʒ/	/ʃ/ mission & fission vs. /ʒ/ vision & fusion
/f v w/	fine vs. vine vs. wine be very wary of Victor Walter Viceroy of Wales
/tʃ dʒ/	/dʒ/ judges go to /tʃ/ church to /dʒ/ jive
/r/	really very roszy /z/

### Morphology and Syntax

Aside from the phonetic differences between the two languages, there are also differences regarding morphology and syntax. One of the major differences between the two languages is word

inflection, which refers to the different forms a word can take to express a grammatical function or attribute, such as tense, mood, and number. Because Norwegian and English differ in several aspects regarding inflection, L2 learners might find this to be one of the most challenging aspects of English. One of the most common grammatical mistakes that Norwegian learners make is subject verb agreement. In English, verbs are conjugated and undergo inflectional changes based on the plurality of the subject, such as “I *read* a book,” “he/she *reads* a book,” “we/they *read* a book.” Unlike English, Norwegian verbs are not conjugated based on the number of nouns. Instead, Norwegian uses the ending “-er” for all persons in the presents tense as shown in Table 7 below.

**Table 7.** A comparison of Norwegian and English present tense verbs.

Norwegian (å lese)	English (to read)
Jeg <b>leser</b>	I <b>read</b>
Du <b>leser</b>	You (singular) <b>read</b>
Han/Hun/Det <b>leser</b>	He/She/It <b>reads</b>
Vi <b>leser</b>	We <b>read</b>
Dere <b>leser</b>	You (plural) <b>read</b>
De <b>leser</b>	They <b>read</b>

Another inflectional difference is concerned with the changes in nominal suffixes, since nouns in Norwegian undergo changes in suffixes (plural suffix) and clitics (definite article) to reflect gender, plurality, and definiteness. At the same time, adjectives in Norwegian must agree with their nouns in terms of gender and number (Gutman & Avanzati, 2013). While there are also plural suffixes in both languages, English does not have the definite article that attaches to the noun as a clitic. In Norwegian, the definite marker is a clitic that is added to the end of nouns and adjectives to mark definiteness. One unique aspect about adjectives in Norwegian is that it can take double definite markings, which does not occur in English, as in the example “*den store bygningen*” vs. “*en stor bygning*” (“the big building” vs. “a big building”). In terms of the genitive “-s,” it is an enclitic that attaches to the noun—just like in English—however, the apostrophe is omitted in Norwegian. Table 8 below shows examples of inflectional changes in Norwegian.

**Table 8.** Examples of different inflectional changes in Norwegian noun conjugation.

	Singular		Plural	
	Indefinite	Definite	Indefinite	Definite
Masculine	<b>en</b> gut (a boy)	<b>gutten</b> (the boy)	<b>gutter</b> (boys)	<b>guttene</b> (the boys)
Feminine	<b>ei</b> jente (a girl)	<b>jenta</b> (the girl)	<b>jenter</b> (girls)	<b>jentene</b> (the girls)
Neuter	<b>et</b> eple (an apple)	<b>eplet</b> (the apple)	<b>epler</b> (apples)	<b>eplene</b> (the apples)

While morphological and inflectional changes in nouns can pose some confusion for L2 learners with Norwegian L1, it usually does not hinder readers from recognizing grammatically correct sentences from incorrect ones. Conjugations, on the other hand, can create more difficulty for lower level English learners, since conjugation mistakes can lead to syntactic errors.

### **Subject-verb agreement**

In English, the sentence structure usually follows the order *subject-verb-object (SVO)*, where the verb always follows the subject. However, in Norwegian the verb always takes the second place in the main clause, even when the sentence starts with adverbials of time or place. For instance, if the sentence does not start with the subject, then the subject follows after the verb. Below are two examples demonstrating the placement of the verb for English and Norwegian:

(1) Tomorrow Alex travels to Paris.

(2) I morgen reiser Alex til Paris. *Tomorrow travels Alex to Paris.*

In the case of negation, there are also differences between the two languages. For instance, in the placement of the negation in subordinate clauses, English maintains the same negation structure in both the main and subordinate clauses. On the other hand, the negation structure changes in Norwegian in the subordinate clause:

(3) He didn't get what he wanted.

(4) He refused to go because he didn't get what he wanted.

(5) Han FIKK IKKE det han ville. *He didn't get what he wanted.*

(6) Han nektet å gå fordi han IKKE FIKK det han ville. *He refused to go because he didn't get what he wanted.*

As a result, low-level English learners might have a hard time determining the correct placement of the verb, when making grammatical judgements regarding sentence structures in English. In order to have a better understanding of the bilinguals' grammar ability, standard measures of grammatical judgement skills will also be included in the current language experiment.

### **Prosodic features**

Unlike English, Norwegian is a pitch accent language with a two-tone system, which can distinguish word meanings by differences in melodies (Kristofersen, 2000, p. 231). The two tonal melodies are classified as accent 1 and accent 2, and the tonal pattern allows speakers to distinguish between two words with the same sound structure. In EN, accent 1 uses low flat pitch in the first syllable, while accent 2 uses a high, sharply falling pitch in the first syllable and a low pitch in the



beginning of the second syllable, followed by an increase in intonation (Kristoferson, 2000, p. 235). Table 9 below demonstrates a Norwegian minimal pair that distinguishes the first and second accents.

**Table 9.** An example of a Norwegian minimal pair with different tonal accents.

<p>A Norwegian minimal pair</p> <p>a. 'aksel<sub>1</sub> 'shoulder'</p> <p>b. 'aksel<sub>2</sub> 'axle'</p>
---

(Wetterlin, A., 2010)

A problem that many Norwegian learners face when speaking English is rising intonation at the end of a phrase, which gives the quality of a question or an incomplete statement for an English listener. It is often the problem that Norwegian speakers tend to misplace stress with function words such as *and*, *a(n)*, *but*, *the*, *than*, *as*, *have*, *was*, which makes the sentence sound unnatural and non-native like (Hakonsen, 2011). One of the main reasons that Norwegian speakers of English (especially low-level speakers) make this mistake is because Norwegians tend to use many rising tones at the end of words. This is similar to what they would do when speaking in their L1, which gives the language a “singing” quality as opposed to English or other Germanic languages (Hakonsen, 2011, p. 7).

Similar to English, Norwegian also has a primary and secondary stress. The primary stress in Norwegian usually falls on the first syllable, which also occurs in English. However, there are also exceptions, especially in loanwords. In loanwords, the stress can fall on the first, second, penultimate or ultimate syllable (e.g, Dresher, 2013). While the stress can only fall on one place in the syllable—the nucleus—it can fall on different syllables in the word. However, one of the biggest problems that L2 learners face is cognates, given that even though they are spelled the same, they are stressed differently. Because of the different stress patterns between the two languages, it is very likely for less proficient English speakers to transfer knowledge from their L1 to their L2, especially for words with more than one syllable. While more proficient L2 speakers tend to avoid these types of language transfer, lower level English speakers are more likely to succumb to pronunciation mistakes regarding stress placement.

To summarize, while English and Norwegian are both Germanic languages, a number of changes have taken place over time that gave distinctive grammatical, phonemic, and prosodic

features to both languages. We predict that high-level proficiency speakers will not be hindered by their L1 knowledge and will successfully distinguish the two phonemic inventories, as well as making correct grammatical judgements in English. In contrast, speakers of lower L2 proficiency are expected to succumb to these phonemic and syntactical differences between the two languages, and have problems sounding native-like when speaking in their L2. Understanding the essential structural differences between English and Norwegian can allow us to better formulate the behavioral measures that need to be included in our experiment, in order to determine their true language proficiency. The next section will explain the predictions and methods used for this language experiment, as well as the motivation behind the current study.

### **Motivation and predictions**

The current study is a replication and extension of the LEAP-Q (e.g., Marian et al., 2007) and aims to examine the relationship between self-rated language proficiency and objectively measured tasks. Evidence from the LEAP-Q study (2007) demonstrated that self-reported language status is indicative of the bilinguals' actual linguistic performances, and is an efficient tool for assessing language profiles in bilinguals and multilinguals. Another goal of the current language experiment is to examine how self-rated accent proficiency patterns with other aspects of self-rated L2 proficiency, by examining lower level phonological and phonemic proficiency among bilingual speakers. Unlike previous studies, the current experiment uses self-rated accent proficiency as a language variable against objectively measured tasks, which is novel in that we do not know the results. The current study will examine several key aspects regarding self-rated language proficiency and accent proficiency:

- (1) Would answers to questions referring to the same underlying aspects of bilingual profile pattern together in the factor analysis?
- (2) How does self-rated accent proficiency pattern with other levels of proficiency?
- (3) Can self-rated language proficiency predict outcomes of the behavioral tasks?
- (4) How does accent proficiency relate to objective measures?

The current study will include more items regarding accent-related questions for both L1 and L2 (i.e., how important it is for bilinguals to maintain their L1 regional accent; how much effort do bilinguals put into sounding more native-like in their L2; how important it is to have a good L2 accent, etc.). Furthermore, a variety of behavioral tasks will be included to test against each category on the self-rated questionnaire, which includes pronunciation, word reading,

vocabulary scale, spelling, sound elision, phonetic memory, grammar and reading comprehension. These tasks provide a more comprehensive reflection of the participants' true language proficiency, compared to previous studies.

Based on the aim of the study and the key questions under investigation from above, we hypothesize that:

- (1) Variables from the self-assessed questionnaire will pattern together with other variables that reflect the common underlying construct through factor analysis (as it did in the LEAP-Q study).
- (2) Accent-related variables will pattern together with speech-related factors and age-related factors (i.e., L2 speech proficiency, L2 learning speech, and L2 age of acquisition).
- (3) Certain aspects of self-rated language proficiency will be able to predict performance on standardized measures.
- (4) Self-rated accent proficiency might not be indicative of the bilinguals' actual language proficiency due to overestimation and lack of feedback.

Because very little research has explicitly used self-assessed accent proficiency to test against linguistic performances from past experiments, we do not know for sure what the result will look like. However, we do not anticipate self-rated accent proficiency to have strong predicative power over the standardized tasks, since none of the tasks explicitly examines accent proficiency. At the same time, participants could overestimate their accent proficiency, since it is more common to receive feedbacks for text-related factors (such as writing and grammar) than speech related factors. In terms of the factor analysis, we predict that the bilinguals' response from the questionnaire should pattern together to reflect different factors that contribute to their overall language proficiency for both L1 and L2. In regard to variables for L2 accent proficiency, we predict that they have the highest chance of being loaded onto L2 speech-related factors and L2 age-related factors. Age of acquisition could be an important factor, since evidence from previous studies (e.g., Delgado et al., 1999) demonstrated a strong correlation between age of exposure/acquisition and standardized performance. In terms of the battery of behavioral tasks, we predict that proficiency regarding L2 speech and L2 age of acquisition could have predicative influence over tests of phonological skills (i.e., YARK, sound elision, and nonsense word

repetition). In terms of tasks that examine text-related skills (i.e., vocabulary, spelling, reading comprehension, and grammar), we predict that L2 text-related factors should serve as a predictor for those aspects of L2 proficiency. Through the current language experiment, we aim to further confirm the internal validity of the self-rated questionnaire through factor analysis, as well as confirming the relationship between self-rated and standardized proficiency measures through multiple regression analysis.

## **Methods**

### **Participants**

Thirty-seven Norwegian English bilinguals between the ages of 18-35 years old participated in this study. Participants had normal to corrected-to-normal hearing and vision and no known language impairments such as dyslexia or stuttering. All participants were native Norwegian speakers with no second home languages, aside from English. Participants varied in their level of education, from first year University students to PhD candidates. The majority of participants were recruited from the University of Agder and nearby areas through online announcements. Upon completion of the entire experiment<sup>1</sup>, participants received a monetary compensation in the form of a gift card. All testing was conducted in English and procedures were approved by the university ethics committee.

### **Materials and procedure**

The entire experiment was divided into two sections: (1) the self-assessed language questionnaire; and (2) the battery of standardized behavioral tasks. The questionnaire was distributed through SurveyXact, an online-based questionnaire platform. Responses from the questionnaire was automatically stored in a designated database after each session. The entire questionnaire takes around 45 minutes to complete and did not need to be supervised by an experimenter. After the completion of the questionnaire, participants who met all the required criteria were contacted to partake in the second section of the language experiment, where they will be individually presented with a battery of objectively measured tasks.

The self-assessed questionnaire was divided into three sections: (1) demographic and language background; (2) self-assessed rating of language proficiency in Norwegian; and (3) self-

---

<sup>1</sup> There are three parts of the language experiment: self-rated questionnaire, battery of behavioral tasks, and eye-tracking.

assessed language proficiency in English. Items from these three sections were aimed at providing a more comprehensive language profile for the bilingual speaker in both their L1 and L2 language proficiency.

The first section explored the speakers' language background and all the languages that the speaker had been exposed to throughout their lifetime. It included items such as country of birth, country of residence, total years of education, and language of education (with an emphasis on the languages of instruction for each level of schooling) among others. Participants also reported the order in which each language was acquired and how dominant they feel about each language. Participants then rated the percentage in which they used each language on a day-to-day basis, with the total tallying up to 100 percent. Furthermore, the participants stated the cultures that they mostly identified with, as well as their preference for each language in terms of reading and speaking. They were also asked to state whether there was any intrusion between the two languages while speaking, and how often these intrusions occurred. Additionally, participants reported the language they used most often when counting, dreaming, expressing emotions, and inner speech. These questions were selected to not only examine the bilingual speakers' language dominance, but also language preference and language attitude for both their L1 and L2.

The next section examined the participants' self-rated usage and proficiency in their first language (L1) Norwegian. Unlike the original LEAP-Q, the current questionnaire was designed to provide a more detailed language proficiency rating that examined the participants' dominance, preference, and language attitude of their L1. Questions regarding language attitude were of special interest in the current questionnaire, given that the original LEAP-Q did not focus on the participants' attitudes towards their mother tongue (in this case the dialect that the speaker used when speaking Norwegian). In the current questionnaire, participants reported the specific dialect that they use while speaking Norwegian, as well as their preference for their own dialect and how much they are willing to modify it when speaking with someone with a different regional dialect. While questions regarding age of acquisition and duration in a L1 country or environment were comparable with the original questionnaire, questions regarding judgments were made independently for speaking, understanding, reading, writing, grammar, vocabulary, and spelling in the current assessment. These items were categorized separately because the experiment has designed a specific behavioral task to test against each of these areas in the bilingual language

profile. Each item was rated on a scale from 0-10, with 0 as “never/not at all” and 10 as “all the time/extremely important.” Furthermore, the current questionnaire changed some of the contributors that influenced the speakers’ language learning and language exposure in their L1. For instance, instead of having “radio” and “language tapes” as contributors, the revised questionnaire switched them with “streaming” and “listening to media” (e.g. podcast and audiobooks), which reflected more up-to-date modes of language learning for speakers when being exposed to their L1.

The last section examined the participants’ self-rated usage and proficiency in their second language (L2) English. Similar to the previous section, this section also examined the participants’ dominance, preference and language attitude, with the exception that all questions were related to their L2. Participants also reported the specific types of accent they have for English (e.g., British/American/other/none in particular). On a scale of 0-10, participants rated their proficiency for speaking, understanding, reading, writing, grammar, vocabulary, and spelling in their L2. Participants assessed each aspect of their L2 proficiency independently, with 0 being “none” and 10 being “perfect.” In terms of accent production, participants rated how much of their L1 accent intrudes during their L2 production, and how frequently others identify them as non-native speakers (of English) based on their foreign (L1) accent while speaking in their L2. In terms of language attitude, participants reported how important it is for them to have a native-like accent, as well as how much effort they put into improving their L2 accent in order to sound more native-like. All of the questions regarding accent production and language attitude were rated on a scale of 0-10, with 0 being “never/none at all” and 10 being “all the time/extremely important”.

### **Description of battery of behavioral tasks**

Upon completion of the first part of the language experiment, the qualified participants were invited to the language lab at the University of Agder to partake in the behavioral tasks. A specially designed recording room was used for the behavioral tasks inside the language lab, where it has sound proof features for optimum voice recording purposes. All of the tasks were administered to the participants through a computer. A headset was required for certain tasks, and an instruction would appear in front of the screen before the starting of each task. Participants were tested individually with a trained experimenter inside the recording room.

Prior to participating in the language experiment, each participant read and completed the consent form and was instructed to leave all electronic items outside of the recording room. Once

inside the room, the participant was seated (in a comfortable position) in front the testing computer and was assigned a unique candidate number for the language experiment that would be automatically stored in the database. There was a headset next to the computer where the participant was asked to wear it for certain tasks during the experiment. The experimenter was seated next to the participant with a second keyboard for scoring and administering the test.

Instructions appeared in front of the screen in the beginning of each task. Instructions included an overview of the task along with instructions regarding which keys to enter when answering each question. The participant was not allowed to ask the experimenter for any assistance during the experiments. If a technical error occurred, the participant would be asked to continue and return to that particular section after the completion of the entire battery of tasks. The average time for completing all eight tasks is approximately 45 minutes. Participants could choose to view their scores at the end of all eight tasks or exit without viewing it. All information regarding the participants would remain anonymous and not be used for any purpose outside of the language experiment.

There were eight separate behavioral tasks for the participants to complete. Each task was designed to examine a particular aspect of the bilingual speakers' language proficiency, such as tests of lower level phonological and phonetic proficiency, vocabulary, spelling, grammar judgement, and reading comprehension. Table 10 below shows an overview of the eight behavioral tasks in the order that was distributed to the participants along with a brief overview for each task:

**Table 10.** List of standardized behavioral tasks and response required.

<b>Behavioral Tasks</b>	<b>Language ability tested</b>	<b>Response required</b>	<b>Approximate Duration</b>
<b>Sentence Repetition</b>	English pronunciation	Memorize and repeat after the sentence as clearly as possible	0-3 minutes
<b>York Assessment of Reading for Comprehension (YARK)</b>	English phonological skills, alphabetical knowledge and word reading	Read out loud the list of words that appear on the screen as accurately as possible	0-3 minutes
<b>British Picture Vocabulary Scale (BPVS)</b>	English vocabulary in non-native speakers	Selecting the correct image that corresponds to the word spoken through the headset	5-10 minutes
<b>Spelling Test</b>	English words and word perception	Typing the correct spelling into the computer after a word is spoken through the headset	3-5 minutes
<b>Sound Elision</b>	English phonological skills	Repeating and forming new words by eliciting particular sounds of a word	0-3 minutes

<b>Nonsense Word Repetition</b>	Low level manipulation and phonetic memory	Listening and repeating what is said through the headset	0-3 minutes
<b>Morpho-Syntax Test (Grammatical error judgement)</b>	Grammar judgement in English sentence structures	Differentiating grammatically correct sentences from incorrect ones	3-5 minutes
<b>Grey Silent Reading Test (GSRT)</b>	Silent reading comprehension and English analytical skills	Read the text and answer accompanying questions as accurately as possible	Maximum 25 minutes

1. **Sentence Repetition**- The sentence (see Appendix C) was selected for English accent recording, where participants would be judged on the accuracy of their pronunciation in their L2.
2. **York Assessment of Reading for Comprehension (YARC)**- The YARC is a diagnostic reading assessment that examined the participants' reading and comprehension skills. This assessment was created for British school children between 4-16 years old, and assesses the language production for L2 speakers of English. There were seven lists of words that varied in word lengths and difficulty, and participants were required to say each word as loudly and clearly as possible.
3. **British Picture Vocabulary Scale (BPVS)**- The BPVS is a standardized test instrument that determines the receptive vocabulary skills of speakers of English. In this task, a set of four pictures would appear on the screen and the participant had to choose the one that corresponds most closely to the word said through the headphone. There were 70 sets of pictures in total. The entries covered a wide range of language levels, including different word classes such as nouns, verbs, and adjectives.
4. **Spelling Test**- The spelling test was created by the experimenters to test the participants' ability in sound/spelling correspondence. Twenty words were created by the experimenters in order to examine the bilingual's knowledge in general transparent sound and nontransparent sound of English (see Appendix C for list of words). The words were recorded and played through the headset one at a time in the same order for each participant. The participants entered the spelling onto the screen one at a time, before a new word is spoken to them.
5. **Sound Elision**- The sound elision test examined the participants' ability in isolating sounds, especially sounds that do not generally occur in Norwegian. A recorded word was played through the headphone to the participant. The participant was first instructed (from the



recording) to repeat after the word, immediately after, they were instructed to say the same word by omitting a particular sound. For example, *say /'li.ɔgs/, now say /'li.ɔgs/ without the //*. There were 18 sets in total (see Appendix C for full list). This task examined the participants' phonological knowledge and ability in eliciting sounds in English.

6. ***Nonsense Word Repetition (NWR)***- The NWR examined low level manipulation and phonetic memory of L1 Norwegian speakers. Participants repeated after a series nonsense words such as /'θra:n.tɪb/, /kɪf.'geɪ.brʌm/, and /'rɛt.kɪp.təʊ.vʊm.təm/. The words were created and recorded by the creators of this task and were played through the headphone. Each word was played only once and participants had only one chance to correctly repeat after it. The nonsense words became progressively longer and more difficult to produce with each new word. There were 22 nonsense words in total (see Appendix C for full list).
7. ***Morpho-Syntax Test (grammatical error judgement)***- This task examined the participants' ability in making correct grammatical judgements in their L2. There was a total of 32 sentences that examined errors in the following four different categories of grammar: word order errors; verb (V2) placement errors; and subject verb agreement errors (divided into two parts for both plural and singular forms). For each category, eight sentences were created by the experimenter—four correct and four incorrect for the participants to determine. For this task. participants were required to answer as quickly and as accurately as possible.
8. ***Grey Silent Reading Test (GSRT)***- The GSRT is a reading comprehension assessment that provided a quick and efficient measure of silent reading comprehension. This task examined the participants' ability in analyzing information in English, as well as their ability in correctly using text features to make information accessible. There were eight short readings in total, and five multiple choice questions for each reading. The questions appeared one at a time with the story directly above it, and there was only one correct answer. After the participants answered all five questions, a new story would appear. The levels of reading increase with each new story, and the time limit set for this task was 25 minutes. However, participants were not required to finish all eight readings if they exceeded the time limit. Speed was not important for this task.

All of the behavioral tasks were designed to test against a specific aspect of the participants' self-rated language proficiency from the self-rated language profile. Tests such as BPVS, YARC, and GSRT were taken from actual assessments made for English L1 school children, while others such as the spelling, sound elision, nonsense word repetition, and grammar judgement were designed specifically with the Norwegian L1 in mind. Tests that were specifically designed for this experiment can be found in Appendix C.

## Results

Thirty-seven Norwegian English bilingual speakers participated in this experiment. All of the participants were born and raised in Norway with only one native language. There were thirty-eight participants initially, however, one subject had been removed from the analysis due to incomplete data. The average number of years of formal education was 13.4 years, and the majority of the participants were university students at the time of the experiment. All of the participants have been exposed to English since childhood, and most use English on a day-to-day basis. Table 11 below shows the results of the language variables from the self-assessed language profile for participants' L1 and L2 language history and proficiency.

**Table 11.** Self-reported language history and proficiency for participants.

Language variables	L1 history		L2 history	
	Mean	Range	Mean	Range
% time speaking (out of 100%)	57.6%	25-80%	38.3%	16-70%
Language intrusion (0-10)				
• L2 intruding into L1	NA	NA	5.8	0-9
• L1 intruding into L2	2.8	0-9	NA	NA
Immersion duration (months)				
• In a country	287.5	228-420	6.5	0-36
• In a family	284.3	222-420	14.5	0-253
• In a school	239.7	148-377	18.8	0-108
Language learning (0-10)				
• From family	9.3	5-10	3.6	0-10
• From friends	7.6	3-10	5.4	0-10
• From reading	6.8	3-10	7.9	3-10
• From school	7.9	1-10	7.6	3-10
• From TV	4.3	0-10	7.7	3-10
• From music	3.7	0-10	7.4	2-10
Extent of language exposure (0-10)				
• Family	9.4	4-10	2.3	0-8
• Friend	8.5	5-10	5	1-10
• Reading	5.5	1-10	8.5	4-10
• School	2.1	0-10	6.5	0-10
• TV	3.8	1-10	8.4	2-10
• Music	3.5	0-10	8.4	4-10
Language proficiency (0-10)				
• Speaking	7.2	6-8	6.3	2-8
• Understanding	7.3	6-8	7.2	5-8
• Reading	7	3-8	7	5-8
• Writing	6.8	3-8	6.1	2-8

• Grammar	6.3	2-8	5.5	2-8
• Vocabulary	6.4	2-8	5.6	2-8
• Spelling	6.4	2-8	5.8	1-8
• Pronunciation	NA	NA	6	1-8
Regional dialect (0-10)				
• How strongly regional is your spoken Norwegian	7.5	1-10	NA	NA
• How frequently can others identify your regional dialect	7.6	0-10	NA	NA
• Importance of having a native Norwegian dialect	6.4	1-9	NA	NA
• Willingness to modify L1 accent	4.3	1-9	NA	NA
Age (years)				
• Begin acquiring L2	NA	NA	6.3	3-10
• Becoming fluent in speaking L2	NA	NA	15.1	9-30
• Begin reading in L2	NA	NA	7.7	5-13
• Becoming fluent in L2 reading	NA	NA	13.9	9-18
Attitude towards foreign accent (0-10)				
• Self-perceived Norwegian accent when speaking L2	NA	NA	3.1	1-9
• How frequently others identify you as a non-native speaker when speaking L2	NA	NA	4.7	0-9
• Importance of having a native-like L2 accent	NA	NA	7	1-9
• Effort in sounding native-like for L2 accent	NA	NA	5.7	1-9
Note. L1= first language; L2= second language. NA= question does not apply				

*Language proficiency and exposure.* Table 11 reveals several significant language variables that were assessed through the questionnaire. Given that this group of bilinguals were L1 speakers of Norwegian with English as their L2, there was a significant difference between the average time participants spent using each language. According to the data, the average time spent using L1 was 57.6%, while the average time spent using L2 was 38.3%. Contrary to expectation, L2 intruded more frequently into bilinguals' L1 than the other way around, even though the average time spent speaking L1 was significantly higher than L2. There was also a clear difference between L1 and L2 acquisition, given that the highest contributors for L1 learning were family and friends, while the highest contributors for L2 learning were reading, school, and television. There was a big contrast between the nature of acquisition, given that L1 was acquired through a more natural environment (i.e., interaction with family and friends), while L2 was acquired through a more academic setting (i.e., school and books). A similar pattern was shown for language

exposure, since the highest factors for language exposure were family and friends for L1. On the other hand, reading, TV, and music were the highest factors for language exposure for L2. In terms of language proficiency, participants rated speaking, understanding, and reading as the highest for both L1 and L2.

*Language attitude.* Results for items related to language attitude (both for L1 regional dialect and accent proficiency for L2) revealed that maintaining a strong regional dialect is important for bilingual speakers, since it received an average rating of 7.5 (all ratings were rated between 0-10). A similar rating of 7.6 was obtained for how frequently others can identify which region of Norway they are from based on their L1 regional dialect. On the other hand, a much lower score of 4.3 was rated for willingness to modify one's regional dialect when speaking to someone with a different regional dialect. When it comes to attitudes towards foreign accent, the average rating for self-perceived Norwegian accent when speaking L2 was 3.1, which was relatively low. On the other hand, a higher rating of 4.7 was obtained for how frequently they think others can identify their foreign accent when speaking in L2. Lastly, participants reported a relatively high rating of 7 for the importance of having a native-like accent when speaking L2, despite of their unwillingness in modifying their L1 accent/dialect.

*Age of acquisition.* The age of acquisition was only applied to L2, given that all participants were born into L1 Norwegian backgrounds. The average age for begin acquiring L2 was 6.3 years old, which was relatively young, since the average age for the participants were between 18-35 years old. The average age for becoming fluent in speaking L2 was 15.1 years old. In terms of reading, the average age for begin reading in L2 was 7.7 years old, and the average age for becoming fluent in reading L2 was 13.9 years old, which was younger than the average age for becoming fluent in speaking. Evidence for age of acquisition suggested that this group of bilinguals were early bilinguals, since all of them acquired their L2 during childhood and became fluent prior to adulthood.

### **Factors analysis**

Factor analysis was used to analyze the variables from the results of the questionnaire, and to show which variables grouped together. Variables with too high correlations ( $>0.8$ ) and variables that do not correlate at all with any others ( $<0.3$ ) were excluded from the factor analysis. The variables that fell within the correct correlation value were submitted to a factor analysis for which both the optional and parallel analysis suggested 8 factors, which explained 62% of the

variance. These factors along with the variables that loaded onto them are shown in Table 12. Positive correlations indicated that the variables and their loading values are going in the same direction, while negative correlations indicated that they are going in opposite directions. For instance, a negative correlation between age of acquisition and L2 proficiency would be defined as the younger the speakers started learning L2, the more proficient they become. Table 12 below demonstrates the factors yielded from the factor analysis.

**Table 12.** Factors yielded from the self-reported language profile.

Factor 1 (RC1): L2 text proficiency	Loading values	Factor 2 (RC7): L1 proficiency	Loading values	Factor 3 (RC4): L2 speech proficiency	Loading values	Factor 4 (RC2): L1 text proficiency	Loading values
L2 Exposure, reading	0.81	L1 Exposure, friends	0.78	L2 Proficiency, pronunciation	0.82	L1 Proficiency, reading	0.8
L2 Learning, friends	0.77	L2 Exposure, music	0.60	L2 Proficiency, speaking	0.75	L1 Proficiency, speaking	0.75
L2 Exposure, friends	0.77	L1 Regional accent	0.59	L2 Proficiency, vocab	0.66	L1 Proficiency, writing	0.65
L2 Proficiency, understanding	0.65	L1 Learning, reading	0.52	L2 Exposure, family	0.48	L1 Proficiency, spelling	0.54
L2 Proficiency, writing	0.60	L1 Proficiency, vocab	0.50	L2 Proficiency understanding	0.39	L1 Proficiency, understanding	0.54
L2 Proficiency, reading	0.60	L1 Importance, dialect	0.47	Age	0.39	L1 Proficiency, grammar	0.54
L2 Learning, reading	0.48	L1 Proficiency, spelling	0.47	L2 Proficiency, grammar	0.38	L1 Months in School	0.5
L2 Proficiency, spelling	0.47	L1 Exposure, reading	0.46	L1 Proficiency, vocab	0.37	L1 Others identify region	0.47
L2 Proficiency, grammar	0.41	L1 Proficiency, gramma	0.42	L2 Months with family	0.33	L2 Age begin acquiring	0.46
L2 Proficiency, speaking	0.39	L2 Learning, TV	0.38	L2 Months in country	0.30	Age	0.37
L2 into L1 Intrusions	0.38	L1 Modify accent	0.36	L1 Importance, dialect	-0.37	L2 Age began reading	0.37
L1 Exposure, family	0.32	L2 Learning, music	0.36	L1 Exposure, family	-0.41	L2 Percent Time used	-0.37
Modify L1 accent	-0.45	Self perceived L1 accent	0.34	L1 Learning, Family	-0.42	L2 Months with family	-0.63
Age	-0.37	L2 Proficiency writing	0.31	L2 Frequency identified	-0.65		
		L2 Effort with accent	-0.57	L1 self-perceived accent	-0.66		
		Years of formal education	-0.51				
		L2 Exposure, school	-0.45				
		L2 Age begin acquiring	-0.35				
		L2 importance accent	-0.34				
		L2 Age begin reading	-0.34				
		L2 Frequency identified	-0.33				
Proportion Variance	0.09	Proportion Variance	0.09	Proportion Variance	0.09	Proportion Variance	0.08
Cumulative Variance	0.09	Cumulative Variance	0.18	Cumulative Variance	0.26	Cumulative Variance	0.34
Factor 5 (RC5): General language proficiency	Loading values	Factor 6 (RC3): L1 learning	Loading values	Factor 7 (RC6): L2 learning speech	Loading values	Factor 8 (RC8): L2 age of acquisition	Loading values
L2 Learning, school	0.69	L1 Learning, music	0.83	L1 into L2 Intrusions	0.73	L2 Age fluent reading	0.82
L1 Learning, Family	0.68	L1 Exposure, music	0.80	L2 into L1 Intrusions	0.68	L2 Age fluent speaking	0.75
L1 Learning, Friends	0.62	L1 Exposure, TV	0.77	L2 Learning, Family	0.67	L2 Age begin acquiring	0.46
L2 Exposure, spelling	0.59	L1 Learning, TV	0.73	L2 Exposure, family	0.58	L2 Age begin reading	0.39
L1 Proficiency, grammar	0.58	L1 Exposure, reading	0.51	L2 Exposure, TV	0.47	L1 Importance, dialect	0.34
L2 Proficiency, grammar	0.53	L1 Learning, school	0.43	L2 Exposure, music	0.46	L2 importance accent	0.34
L2 Exposure, school	0.51	L1 Learning, reading	0.42	L2 Percent Time used	0.46	L2 Learning, school	0.33
L2 Months in School/Workplace	0.48	Years of formal education	0.42	L2 importance accent	0.44	L2 Proficiency grammar	-0.31
L1 Proficiency, spelling	0.44	L2 Learning, reading	0.36	L2 Learning, music	0.43	L1 Learning, school	-0.39
L1 Proficiency, writing	0.43	L2 Learning, TV	0.34	L2 Learning, TV	0.35	L1 Proficiency, understanding	-0.49
L1 Exposure, family	0.39	L2 Learning, music	0.34	L2 Age began reading	0.32		
L1 Learning, Reading	0.37	L1 Modify accent	-0.4	L2 Proficiency reading	0.32		
L2 Months in country	0.36						
L1 Proficiency, vocab	0.35						
L2 Proficiency, writing	0.31						
L2 Age began reading	0.31						
L2 Months with family	-0.4						
Proportion Variance	0.08	Proportion Variance	0.07	Proportion Variance	0.06	Proportion Variance	0.06
Cumulative Variance	0.42	Cumulative Variance	0.50	Cumulative Variance	0.56	Cumulative Variance	0.62

The first factor had variables that mainly reflected the bilinguals' L2 proficiency. The variables with the highest loading values included L2 exposure through reading, L2 learning through friends, and L2 exposure through friends. It also included variables relating to L2 proficiency in understanding, writing, reading, spelling, grammar, and speaking. There was also one variable relating to L1 exposure, and two negatively loading variables relating to L1 accent and age. Together the variables under this factor revealed an underlying construct of L2 text proficiency.

The second factor had variables that mainly reflected the bilinguals' L1 proficiency, with L1 related variables loading positively onto this factor. The variables with the highest loadings included L1 exposure through friends and music, L1 regional accent, L1 learning through reading, L1 proficiency in vocabulary, spelling, and grammar, as well as L1 exposure through reading. There were also several L2 variables that loaded negatively onto this factor. Some of those variables included effort speakers put into sounding more native-like in their L2, years of formal education, L2 exposure from school, and importance of having a good L2 accent among others. Together, the positive loadings for L1 variables along with the negative loadings for the L2 variables reflected an underlying construct of L1 proficiency.

The third factor had variables that mainly related to L2 speech proficiency, which all had positive loadings, except for one (i.e., *how frequently do others identify your foreign accent when speaking in L2*). The variables with the highest loadings included L2 proficiency in pronunciation, speaking, and vocabulary. It also included L2 exposure to family, and L2 proficiency in understanding, grammar, and vocabulary. There were also several variables with negative loadings regarding the bilinguals' L1. These variables mainly reflected the bilinguals' language attitude and preference, which included the importance of maintaining their L1 dialect, L1 exposure to family, L1 learning from family, and their self-perceived L1 accent when speaking in their L2. Given that the variables that loaded onto this factor were mainly associated with different aspects of L2 speech, this factor can be interpreted as L2 speech proficiency.

The fourth factor had variables mainly reflecting L1 text-related proficiency. The variables with the highest loading values included L1 proficiency for reading, speaking, writing, spelling, vocabulary, and grammar. It also included the amount of time the speakers spent in school where their L1 was used. In addition, there were two L2 variables that loaded positively onto this factor, and two L2 variables that loaded negatively onto this factor. The two negatively loading variables included percent of time spent using L2 and number of months spent with family where L2 was spoken. The positive loadings relating to L1 along with the negative loadings relating to L2 indicated that this factor has an underlying representation of L1 text proficiency.

The fifth factor included almost an equal number of variables relating to both L1 and L2 for language exposure and proficiency. Among the variables, the highest loadings included L2 learning through school, L1 learning through family, L2 exposure through spelling, and

proficiency in grammar for both L1 and L2. The variables under this factor mainly reflected L1 and L2 exposure and proficiency, with only one negatively loading variable in regard to the amount of time spent in a L2 environment with family. As a result, the variables loaded under this factor (both L1 and L2) revealed an underlying construct of general language proficiency.

The sixth factor had variables that mainly reflected L1 learning and exposure. The variables with the highest loading values included L1 learning through music, L1 exposure through music and TV, as well as L1 learning through TV, among others. There were also three L2 related variables that loaded positively onto this factor, all relating to different forms of L2 learning (i.e., reading, TV, and music). There was only one negatively loading variable relating to how much the speakers would modify their L1 dialect when conversing with someone with a different dialect. Through the variables that loaded together onto this factor, evidence suggests that this factor reflected L1 learning.

The seventh factor had variables that mainly reflecting L2 learning for speech. All of the variables were related to L2 and loaded positively onto this factor. The variables with the highest loading values included language intrusion (both from L1 into L2 and L2 into L1), L2 learning and exposure, L2 percent time used, importance of having a native-like L2 accent, and L2 age-related variables. Given that all of the positively loading factors were related to different aspects of L2 learning, especially when it comes to speech production, evidence suggests that the underlying representation for this factor is L2 learning speech.

The eighth factor included variables that mainly reflected L2 age of acquisition. The variables with the highest loading values included L2 age for becoming fluent in reading, L2 age for becoming fluent in speaking, as well as L2 age for begin acquiring, and L2 age for begin reading. There were also three negatively loading variables relating to L2 proficiency in grammar, L1 learning from school, and L1 proficiency in understanding. The positive loadings for L2 age-related variables along with the negative loadings for L1 learning and proficiency related variables reflected an underlying representation of L2 age of acquisition.

### **Correlations between behavioral and self-reported measures**

Similar to the LEAP-Q study conducted by Marian, Blumenfeld, and Kaushankaya (2007), the present study also used multiple regression analyses to generate predictive relationships between self-reported proficiency and objectively measured proficiency. Table 13 below demonstrates the outcomes of the behavioral tasks, and numbers with an asterisk on the side falls

within the significance value  $*p < .05$ , thus serving as a predictor of the measurement. The only task that was excluded from the multiple regression analysis was sentence repetition.

**Table 13.** Result of multiple regression analyses

Measures	Regression coefficients				Predictor
Grey Silent Reading Test (GSRT)	0.66649	0.21462	2.125	0.04*	(RC2): L1 Text proficiency
	0.62295	0.31280	1.922	0.06	(RC4): L2 Speech proficiency
British Picture Vocabulary Scale (BPVS)	0.88490	0.26594	3.327	0.00*	(RC2): L1 Text proficiency
	0.77138	0.3629	2.518	0.02*	(RC5): General language proficiency
	0.88208	0.27167	3.247	0.00*	(RC3): L1 Learning
	-0.65893	0.34884	-1.889	0.07	(RC8): L2 Age of acquisition
York Assessment of Reading for Comprehension (YARK)	0.63164	0.28193	2.240	0.03*	(RC5): General language proficiency
Spelling	0.48453	0.24479	1.979	0.06	(RC2): L1 Text proficiency
	1.3341	0.7155	1.864	0.07	(RC5): General language proficiency
Grammar judgement (reaction time)	-1.2506	0.6347	-1.971	0.06	(RC3): L1 Learning
	1220.6	571.9	2.134	0.04*	(RC8): L2 Age of acquisition
Grammar judgement (%correct)	-793.0	416.2	-1.905	0.07	(RC7): L1 Proficiency
	0.710498	0.37144	1.913	0.07	(RC5): General language proficiency
Sound Elision	0.86779	0.41421	2.095	0.05*	(RC5): General language proficiency
	-1.09164	0.47176	-2.314	0.03*	(RC8): L2 age of acquisition
Nonsense word repetition	No significant effect				

The results yielded positive correlations between standardized tests (i.e., reading comprehension, vocabulary knowledge, phonological skills, and grammatical judgments) and self-reported measures of L1 proficiency, general language proficiency, and L2 age of acquisition. Neither L2 proficiency for text nor speech played a significant role in predicting the outcome of the standardized tests. The majority of the standardized measures correlated more strongly with self-reported general language proficiency, rather than L2 proficiency alone. For L1, self-reported proficiency measures correlated most strongly with reading comprehension (GSRT) and vocabulary knowledge (YARK). For L2, the highest correlation values were generated for grammar judgement (reaction time) and sound elision. Unlike the results obtained by Marian and colleagues (2007), self-rated L2 proficiency did not play a predicative role for reading comprehension and oral production. Below is a brief explanation for the result of each task underneath.

**BPVS.** Table 13 revealed four predictors that emerged from the analysis for the vocabulary knowledge task. Out of the four variables, only three fell within the significance value. The three



variables with significance influence were: L1 text proficiency, general language proficiency, and L1 learning.

**YARK.** Two predictors were generated from the multiple regression analyses, which were general language proficiency and L1 text proficiency. General language proficiency was the only variable that fell within the significance value.

**Spelling.** Two predictors were generated for this task (i.e., general language learning and L1 learning), but neither of the two fell within the significance value.

**Grammar judgement.** Data for the grammar judgement knowledge was divided into two sections: reaction time and percentage of accuracy. For reaction time, the only factor that fell under the significance value was L2 age of acquisition. In term of percentage correct, general language proficiency was the only predictor generated for this task. However, it did not fall within the significance value.

**Sound elision.** The two factors that served as the highest predictors for this task were general language proficiency and L2 age of acquisition. Both factors fell under the significance value for this task.

**NWR.** No predicative relationships were generated for nonsense word repetition from the multiple regression analyses.

The next section will discuss the findings from both the factor analysis and multiple regression analyses more closely, especially by comparing the initial hypotheses made prior to the language experiment and the results obtained through the standardized measurements. Finally, I will discuss the role of L2 speech proficiency and self-rated accent proficiency, and provide explanations for why they failed to predict any of the behavioral measures.

## **Discussion**

The aim of the current study was to investigate the bilingual processing of proficient Norwegian-English speakers by examining the relationship of self-rated L2 proficiency and linguistic performances from standardized measures. This study also examined self-rated accent proficiency and how it relates to other aspects of L2 proficiency, which has not been explored in previous studies. Prior to the language experiment, we predicted that results from the self-assessed questionnaire will pattern together with other variables with the common underlying construct through factor analysis, and accent proficiency will most likely pattern with L2 speech production

and L2 age of acquisition. We further predicted that some factors from the factor analysis (i.e., L2 speech proficiency and L2 text proficiency) can serve as predictors for the behavioral tasks, but self-rated accent proficiency will not have a strong predicative influence. First, we will discuss some of the key findings obtained from the experiment, before discussing results from the factor analysis and multiple regression analyses.

Based on the results obtained from the study, there are some distinctive patterns that separated the bilinguals' L1 and L2, especially for language learning (i.e., variables for how each language was acquired), language exposure, and language intrusion. Despite of having similar ratings for language proficiency in both L1 and L2, evidence suggested that self-rated language proficiency for L1 was higher than L2 across all aspects of language proficiency. The factor analysis was used to analyze how the variables patterned together to form underlying constructs of the bilinguals' language profile. While variables reflecting the common underlying construct did pattern together (see Table 12 for reference), variables related to L2 accent proficiency did not load onto all L2-related variables. Below, we will discuss how self-rated accent proficiency patterned with other levels of L2 proficiency through the factor analysis.

The first factor reflected L2 text proficiency, and the only accent related variable under it was how willing the bilinguals are in modifying their L1 accent when speaking with someone from another region (of Norway). This variable loaded negatively onto the factor, which indicated that L1 accent-related variables did not contribute to L2 text proficiency. No L2 accent-related variables loaded onto this factor, which suggested L2 accent proficiency was not reflected through L2 text proficiency.

The second factor reflected L1 proficiency, and several accent-related variables uploaded onto this factor. Questions regarding the bilinguals' L1 regional accent, importance in maintaining their L1 dialect, degree of willingness in modifying their L1 accent, and self-perceived L1 accent when speaking L2 all loaded positively onto this factor. This indicated that L1 proficiency was reflected through their self-assessed L1 accent proficiency. On the other hand, L2 accent-related variables all loaded negatively onto this factor. These variables included effort put into improving L2 accent, importance of having a good L2 accent, and how frequently others can detect their foreign accent when speaking L2. The negative loadings for L2 accent-related variables indicated that the higher

the L1 proficiency, the lower the L2 accent-related proficiency, since there was a negative correlation between the two.

The third factor reflected L2 speech proficiency, and there were three variables related to accent proficiency, and all three loaded negatively onto the variable. The first one was the importance of maintaining L1 dialect, the second one was how frequently others can detect their foreign accent when speaking in L2, and the last one was self-perceived L1 accent when speaking L2. This indicated that as the bilinguals' speech proficiency increases, their frequency of foreign accent being detected decreases, and their self-perceived L1 accent during L2 speech production also decreases. While L2 accent proficiency does not explicitly get reflected under this factor, evidence does suggest that an increase in L2 speech proficiency leads to a decrease in foreign accent during L2 speech production.

The fourth factor reflected L1 text proficiency, and no variables relating to self-rated accent proficiency loaded onto this factor. This indicated that neither L1 nor L2 accent proficiency was reflected through L1 text proficiency, since it does not involve speech production.

The fifth factor reflected general language proficiency, and similar to Factor 4, no variables relating to self-rated accent proficiency loaded onto this factor. Even though variables under this factor reflected proficiency in spelling, grammar, reading, vocabulary, and writing for both L1 and L2, accent proficiency did not get reflected. This is interesting since speaking should play a big role in the bilinguals' general language proficiency, yet no variables related to accent or speech appeared under this variable. In other words, the bilinguals' general language proficiency did not reflect their speech or accent proficiency.

The sixth factor reflected L1 learning, and the only accent-related variable was the bilinguals' willingness in modifying their L1 accent when speaking to someone with a different accent/regional dialect. There was a negative correlation between willingness in modifying L1 accent and L1 learning, which suggested that the longer they are exposed to their L1 (especially to a certain dialect), the less willing they are to modify it. This indirectly reflected the bilinguals' adherence and preference to their own L1 dialect, which could influence their production of L2. This was also reflected on the self-rated questionnaire, since bilinguals only rated a 4.3 (out of 10) for their willingness in modifying their L1 dialect.

The seventh factor reflected L2 learning for speech, and only one accent-related L2 variable (i.e., importance of having a good L2 accent) loaded positively onto it. This indicated that the more proficient the bilingual speakers become in their L2 speech, the more important it is for them to have a good L2 accent. While this variable reflected more of the bilinguals' language attitude than language proficiency, it does suggest that the bilingual speakers care more about their L2 accent as they become more proficient in L2 speech production.

The last factor reflected L2 age of acquisition, and two accent-related variables loaded positively onto it. The first one was importance in maintaining L1 dialect, and the second one was importance in having a good L2 dialect. This suggested that as the bilinguals get older, they also pay more attention to their accent production, and have a stronger desire in sounding more native-like. One explanation could be that children are not as aware of the way they "sound" as adults or learners who start acquiring L2 at an older age. As a result, when bilinguals started acquiring L2 later in life, they are also more aware of the distinctions between their own speech and the speech of others, which explains why it is more important for them to have a good L2 accent.

Based on the results from the factor analysis, evidence suggests that variables that reflect a common underlying construct did pattern together, which confirmed our hypothesis. In addition, self-rated accent proficiency did load onto other aspects of L2 proficiency, such as L2 speech proficiency, L2 learning speech, and L2 age of acquisition. However, it was not reflected under the bilinguals' general language proficiency, and proficiencies regarding L1 and L2 text. Nevertheless, based on the results from the factor analysis, we can conclude that the bilinguals' accent proficiency is directly related to their L2 speech proficiency and L2 age of acquisition.

Multiple regression analyses were used to generate predicative relationships between the self-rated questionnaire and standardized language performances for the bilingual speakers' L2 proficiency. Results obtained from the multiple regression analyses demonstrated that this group of bilinguals were not balanced bilinguals, since their L1 linguistic proficiency had more predicative power over the standardized measurements than their L2. According to the analysis, the strongest indicator for reading comprehension (GSRT) and vocabulary knowledge (BPVS) was L1 text proficiency. Aside from L1 text proficiency, general language proficiency and L1 learning were also predictors of the vocabulary test. This suggested that the participants were more dependent on their L1 when it comes to analyzing texts and vocabulary words. Evidence from the

RHM suggests that asymmetric bilinguals who are not very proficient in their L2 usually depend more on their L1 to get access to respective meanings of the information (e.g., Kroll, 1994; Sunderman & Kroll, 2006). The fact that L2-related factors did not play a role in these two tasks suggested that the bilinguals' self-rated L2 language proficiency was not indicative of their actual linguistic performance. Furthermore, accent proficiency also did not serve any predicative roles in any of the objectively measured tasks, despite of the relatively high self-ratings from the language profile.

In terms of results obtained for the oral reading task (YARC), general language proficiency was the strongest predictor. The general language proficiency factor was made up of different aspects of self-rated language exposure and language proficiency for both L1 and L2. The fact that this factor was the highest predictor indicated that the bilinguals' self-rated general language proficiency did match their performance for this task. Interestingly, none of the variables under this factor was speech-related, even though this task tested the bilinguals' oral production. One interpretation could be that there is a positive correlation between the amount of language learning/exposure and the bilinguals' speech production. In other words, the longer the participants studied and lived in a L2 speaking country, the better they are at L2 oral production.

Results for the spelling and grammar judgement (percent of correct responses) revealed that none of the factors were significant enough to predict the outcomes of these two tasks. The average ratings for these two items under the self-rated questionnaire were the lowest among the different aspects of language proficiency (for both L1 and L2). This could indicate that this group of participants were not particularly confident about their language skills in these two areas for both languages. Even though the self-rated scores for these two subjects were the lowest in comparison to other aspects of language proficiency, they were still on the higher end of the rating spectrum (see Table 11 for reference). This could indicate that even though the participants were correct in reflecting their weakest areas in their language proficiency, they still overestimated their true levels of proficiency for both spelling and grammar. In terms of reaction time for grammar judgement, the strongest predictor was L2 age of acquisition. This may suggest that the younger the bilinguals started acquiring L2, the better they are at judging grammatical mistakes.

For the sound elision task, general language proficiency and L2 age of acquisition were the strongest indicators. The nature of the sound elision task tested the participants' phonological skills,

which was reflected through their self-rated general language proficiency. Variables under the general language proficiency demonstrated the participants' self-ratings for their overall L1 and L2 language proficiency (except for proficiency regarding speech), which was reflected through the sound elision task. In other words, the participants' self-ratings for their general language proficiency did have predicative effects on certain aspects of their L2 proficiency. At the same time, L2 age of acquisition was another predictor for this task. According to Delgado (e.g., Delgado et al., 1999), early exposure may have provided the bilinguals with sufficient practice in speaking, which could be reflected through the sound elision task. For instance, data from Table 11 reveals that the average age for L2 acquisition is 6.3 years old, and the average age for becoming fluent in speaking is 15.1 years, which indicated that this group of speakers were early bilinguals. This further suggested that there is a strong correlation between age of acquisition and language production.

Even though the results from the multiple regression analyses confirmed part of our hypothesis regarding the predicative relationship between self-rated language proficiency and standardized tests, none of the predictors were related to L2 speech or L2 text proficiency. There are several explanations for why self-rated L2 proficiency did not prove to be a very reliable source of language judgement, especially when it comes to speech production. While it is customary for L2 learners to receive direct feedbacks for text-related and grammar-related tasks, they rarely receive direct feedbacks for the way they speak. Furthermore, studies have found that bilinguals tend to overestimate their skill in the weaker language (e.g, Orsmond, Merry, & Reiling, 1997), which was also observed from the current study, where their ratings for L2 proficiency did not match that of the standardized measurements.

While past studies (e.g., Chakarborty et al., 2011; De Jong et al. 2012; De Jong et al., 2013) have shown a correlation between L2 speech proficiency/fluency and L2 linguistic skills (i.e., linguistic knowledge, linguistic processing, and pronunciation), the current study could not confirm such correlations. Unlike previous studies, none of the behavioral tests in the current study served as a measurement of accent production, or took speech fluency into consideration. Instead we only examined the bilinguals' self-ratings. Furthermore, one can argue that there is no objective way of measuring the bilinguals' accent proficiency, since there are also different varieties of dialect in English (i.e., American English, British English, Indian English, etc.). While participants

stated which form of English they speak, the results were purely self-assessed, which could differ drastically from their actual speech production. Instead, we should take speech fluency into consideration for future studies, given that just the phonological aspects of production alone cannot produce an accurate reflection of the bilinguals' L2 speech proficiency, as observed in previous bilingual language studies (e.g., De Jong et al. 2012; De Jong et al., 2013).

For future studies, more tasks examining speech fluency should be included in the battery of behavioral tasks. Furthermore, more questions regarding the criteria in which the bilinguals use to judge their own language proficiency (i.e., feedbacks from teachers, peers, family, etc.) should also be included in the questionnaire. By having a better understanding of how the bilinguals rate their own language skills (especially when it comes to oral proficiency), could help us determine the credibility of their self-assessments. At the same time, the current study was carried out by two experimenters of different L1 backgrounds, and certain tasks (i.e., YARC, sound elision, and nonsense word repetition) required the experimenter's personal judgement to determine whether the participants' oral production was correct. This could lead to a small degree of discrepancy in the evaluation of the bilinguals' behavioral performances, especially for speech production. Therefore, prior to (or after the completion of) the actual language experiments, an inter-rater reliability test should be carried out, to prevent any unnecessary discrepancy.

## **Conclusion**

In conclusion, while the self-ratings from the language profile did not reflect L2 proficiency through the standardized tests, it did provide a general pattern that reflected the bilinguals' L1 and L2 language acquisition and language experience. Questions regarding language acquisition, exposure, and usage did provide information regarding the general nature of the current bilinguals, and showed distinctive underlying constructs that contributed to their L2 learning. For instance, family and friends were the biggest contributors for L1 learning, while reading, school, and media-related variables (i.e., TV and music) were the biggest contributors for L2 learning. At the same time, the data revealed that the participants rated speaking, understanding, and reading as the highest aspects of language proficiency for both L1 and L2, which suggested that they were aware of their strengths and weaknesses in their L2. Even though their self-rated L2 proficiency was not reflected through any of the standardized tests, their self-rated general language proficiency did predict the outcomes for three of the behavioral tasks, and L2 age of acquisition was also a predictor for two of the tasks. This indicated that self-rated language

proficiency did reflect certain aspects of their actual proficiency, and provided general information regarding the bilinguals' language usage and experience.

As predicted in our hypothesis, self-rated accent proficiency did not carry any predicative power over the behavioral tasks. However, it did pattern together with other levels of language proficiency in the factor analysis, such as L2 speech proficiency, L2 learning speech, and L2 age of acquisition. Furthermore, there is a positive correlation between L2 age of acquisition and importance of having a good L2 accent, given that as the bilinguals get older, they also become more aware of their own speech production. As the first study that explicitly examined the relationship between self-rated accent proficiency and how it relates to other aspects of L2 proficiency, data from this language experiment offered further insight into bilingual language processing. In addition, the current language study also provided a more comprehensive measurement of the bilinguals' overall L2 proficiency, both in terms of the items under the self-assessed language profile, and the battery of standardized tasks. The current questionnaire is a more comprehensive tool for assessing L2 language proficiency compared to previous language profiles, and captured factors that could predict linguistic performance through standardized language measurements.



## References

- Anderson-Hsieh, Janet & Johnson, Ruth & Koehler, K. (2006). The Relationship Between Native Speaker Judgments of Nonnative Pronunciation and Deviance in Segmentals, Prosody, and Syllable Structure. *Language Learning*, 42, 529 - 555. doi:10.1111/j.1467-1770.1992.tb01043.x
- Bialystok, E., Craik, F.I., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in Cognitive Sciences*, 16(4), 240-50. doi:10.1016/j.tics.2012.03.001
- Chakraborty, R., Domsch, C., & Gonzales, M.D. (2011). Articulatory behaviors of nonnative speakers: Role of L2 proficiency and accent modification. *Perceptual and motor skills*, 113(1), 311-30. doi:10.2466/11.17.21.28.PMS.113.4.311-330.
- Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(5), 1283-1296. doi:10.1037/0278-7393.26.5.1283
- Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, 50(4), 491-511. doi:10.1016/j.jml.2004.02.002
- De Jong, N.H., Groenhout, H., Schoonen, R., Hulstijn, J.H. (2013). Second language fluency: speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics*, 36(2), 223-243. doi:10.1017/S0142716413000210
- De Jong, N.H., Steinel, M., Florijn, A., Schoonen, R., & Hulstijn, J.H. (2012). Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics*. 34(5). doi:10.1017/S0142716412000069.
- Delgado, P., Guerrero, G., Goggin, J., & Ellis, B. (1999, February). Self-Assessment of Linguistic Skills by Bilingual Hispanics. *Hispanic Journal of Behavioral Sciences*, 31-46.
- Dell, G.S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93(3), 283-321. doi:10.1037/0033-295X.93.3.283
- Derwing, T.M., & Munro, M.J. (1997). Accent, intelligibility, and comprehensibility: evidence from four L1s. *Studies in Second Language Acquisition*, 19(1), 1-16. Retrieved from [https://www.researchgate.net/publication/279568366\\_Accent\\_Intelligibility\\_and\\_Comprehensibility\\_Evidence\\_from\\_Four\\_L1s](https://www.researchgate.net/publication/279568366_Accent_Intelligibility_and_Comprehensibility_Evidence_from_Four_L1s)
- Derwing, T.M., Munro, M., Thomson, R.I., & Rossiter, M.J. (2009). The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition*, 31, 533-557. doi:10.1017/S0272263109990015.

- Desmet, T., & Duyck, W. (2007). Bilingual language processing. *Language and Linguistic Compass*, 1(3), 168-194. doi:10.1111/j.1749-818x.2007.00008.x
- Dollaghan, C., & Campbell, T.F. (1998). Nonword repetition and child language impairment. *Journal of speech, language, hearing, and research*, 41(5), 1136-46. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/9771635>
- Dresher, B.E. (2013). The influence of loanwords on Norwegian and English stress. *Nordlyd*, 40(1), 43-55. Retrieved from <https://septentrio.uit.no/index.php/nordlyd/article/view/2475/2332>
- Fledge, J.M., & Dividian, R.D. (1984). Transfer and developmental processes in adult foreign language speech production. *Applied Psycholinguistics*, 5, 323-347. Retrieved from [http://jimflege.com/files/Flege\\_Davidian\\_transfer\\_processes\\_AP\\_1984.pdf](http://jimflege.com/files/Flege_Davidian_transfer_processes_AP_1984.pdf)
- Gathercole, S.E., & Baddeley, A.D. (1990). Phonological memory deficits in language disordered children: is there a causal connection? *Journal of memory and language*, 29, 336-360. Retrieved from <https://pdfs.semanticscholar.org/e7fc/2a7e9778e2bcece8165e0c484e280812f81c.pdf>
- Gollan, T.H., & Acenas, L.A. (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(1), 246-269. doi:10.1037/0278-7393.30.1.246
- Gollan, T.H., Forster, K.I., & Frost, R. (1997). Translation priming with different scripts: masked priming with cognates and noncognates in Hebrew-English bilinguals. *Journal of Experimental Psychology: Learning Memory and Cognition*, 25(5), 1122-1139. Retrieved from <https://pdfs.semanticscholar.org/299e/21a83c1cda98ce089874c9c6eb613219fe5a.pdf>
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1, 67-81.
- Hakonsen, H. (2011). *Common problems with English pronunciation by Norwegian learners*. (Master's Thesis, Hawaii Pacific University). Retrieved from [https://hannehakonsen.weebly.com/uploads/2/2/0/9/22093434/phonology\\_research\\_paper\\_pdf2.pdf](https://hannehakonsen.weebly.com/uploads/2/2/0/9/22093434/phonology_research_paper_pdf2.pdf)
- Hoffmann, C. (1991). *An introduction to bilingualism*. London; New York: Longman.
- Hoshino, N., Kroll, J.F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition*, 106(1), 501-511. Retrieved from <https://doi.org/10.1016/j.cognition.2007.02.001>

- Kristoffersen, G. (2000). *The Phonology of Norwegian*. Oxford University Press. Retrieved from <http://www.jstor.org/stable/4420207>
- Koster, C.J., Koet, T. (2006). The evaluation of accent in the English of Dutchmen. *Language Learnin*, 43(1), 69-92. doi: 10.1111/j.1467-1770.1993.tb00173.x
- Kroll, J.F. (2008). Juggling two languages in one mind. *Psychological Science Agenda*, 22(1), 1-6. Retrieved from [https://www.researchgate.net/publication/265566523\\_Juggling\\_Two\\_Languages\\_in\\_One\\_Mind](https://www.researchgate.net/publication/265566523_Juggling_Two_Languages_in_One_Mind)
- Kroll, J. F., & Curley, J. (1988). Lexical memory in novice bilinguals: the role of concepts in retrieving second language words. *Practical Aspects of Memory*, 2, 389-395. Retrieved from [https://www.scirp.org/\(S\(351jmbntvnsjt1aadkpozje\)\)/reference/ReferencesPapers.aspx?ReferenceID=1749442](https://www.scirp.org/(S(351jmbntvnsjt1aadkpozje))/reference/ReferencesPapers.aspx?ReferenceID=1749442) (\*\*not sure if this is where I retrieved the article)
- Kroll, J.F., Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations.
- Kroll, J.F., Tokowicz, N. (2005). Models of bilingual representation and processing: looking back and into the future.
- Marian, V., Blumenfeld, H., & Kaushanskaya, M. (2007). *The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals*. Evanston: American Speech-Language-Hearing Association.
- McClelland, J., & Rumelhart, D. (1981). An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account of Basic Findings. *Psychological Review*.
- Meuter, R.F.I., Allport, A. (1999). Bilingual language switching in naming: asymmetrical costs of language selection. *Journal of Memory and Language*, 40, 25-40. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.568.7988&rep=rep1&type=pdf>
- Munro, M.J., Derwing, T.M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 73-97. Retrieved from <https://doi.org/10.1111/j.1467-1770.1995.tb00963.x>
- Poort, E.D., Rodd, J.M. (2017). The cognate facilitation effect in bilingual lexical decision is influence by stimulus list composition. *Acta Psychologica*, 180, 52-63. doi: 10.1016/j.actpsy.2017.08.008
- Segalowitz, N. (2016). Second language fluency and its underlying cognitive and social determinants. *International Review of Applied Linguistics*, 2016, 54(2), 79-95. doi:10.1515/iral-2016-9991

- Shriberg, L.D., Kwiatkowski, J., & Gruber, F.A. (1994). Developmental phonological disorders II: short-term speech-sound normalization. *Journal of Speech and Hearing Research*, 37, 1127-1150. Retrieved from <https://pdfs.semanticscholar.org/a961/deef8cbd807d69aecc35d66e5366832dd6ba.pdf>
- Sunderman, G., & Kroll, J.F. (2006). First Language Activation During Second Language Lexical Processing. *Studies in Second Language Acquisition*, 28, 387-422.
- Talamas, Kroll, & Dufour. (1999). Form related errors in second language learning: A preliminary stage in the acquisition of L2 vocabulary. *Bilingualism: Language and Cognition*, 45-58.
- Van Heuven, W., Dijkstra, T., & Grainger, J. (1998). *Orthographic Neighborhood Effects in Bilingual Word Recognition*. Academic Press.
- Waters, G.S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *The Quarterly Journal of Experimental Psychology. A, Human experimental psychology*, 49(1), 51-57. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/8920099>
- Wetterlin, A. (2010). *Tonal Accents in Norwegian: Phonology, morphology and lexical specification*. [Linguistische Arbeiten 535] Berlin: de Gruyter. pg. 3.
- Wetterlin, A. (2017). EN-122 English Phonetics, 20.09.2017 Class 4 (RP phoneme inventory - comparison) [PowerPoint slides]. Kristiansand.
- Woodcock, R., & Muñoz-Sandoval, A. (1993). *Woodcock-Muñoz Language Survey Comprehensive*. Chicago: Riverside.

## Appendix A

### Revised self-assessed language questionnaire

Are you a native speaker of Norwegian?

- (1)  Yes  
(2)  No

Is Norwegian the only language you speak at home (aside from perhaps English)?

- (1)  Yes  
(2)  No

Are you a reasonably good speaker of English?

- (1)  Yes  
(2)  No

Do you have normal vision or vision that is corrected to normal with glasses or lenses?

- (1)  Yes  
(2)  No

Do you have normal or corrected to normal hearing?

- (1)  Yes  
(2)  No

Can you confirm that you have no language impairments such as dyslexia, stuttering etc?

- (1)  Yes  
(2)  No

### Section 1- Background information.

Full name?

---

---

Age?

---

E-mail address?

---

Date of birth? (DDMMYY)

---

Gender?

- (1)  Male  
(2)  Female

Are you left- or right handed?

- (1)  Left  
(2)  Right

Country of birth?

---

Country of residence?

---

How many years of formal education do you have?

---

What is the highest level of education you have achieved or its approximate equivalent? (e.g. high school, Bachelor, Masters, professional training, PhD, M.D. etc.)

---

**Please write down the name(s) of the language(s) in which you received instruction in school, for each level of schooling:**

Primary/elementary school \_\_\_\_\_

Secondary/middle school \_\_\_\_\_

High school \_\_\_\_\_

College/university \_\_\_\_\_

**Please list all the languages you have learned in the order you first (learned/acquired) them.**

Language 1 \_\_\_\_\_

Language 2 \_\_\_\_\_

Language 2 \_\_\_\_\_

Language 4 \_\_\_\_\_

Language 5 \_\_\_\_\_

**Please list all the languages you can speak in the order of how dominant they are for you.**

Language 1 \_\_\_\_\_

Language 2 \_\_\_\_\_

Language 2 \_\_\_\_\_

Language 4 \_\_\_\_\_

Language 5 \_\_\_\_\_

**Please list what percentage of time you are on average exposed to each language (e.g. exposure in terms of talking, listening, and reading, including TV, films, music, etc.) Percentages should add up to 100.**

Language

1 \_\_\_\_\_ %

2 \_\_\_\_\_ %

3 \_\_\_\_\_ %

Language

4 \_\_\_\_\_ %

5 \_\_\_\_\_ %

**Which language do you prefer to READ in?**  
Please list your languages in order of preference for reading.

Language 1 \_\_\_\_\_

Language 2 \_\_\_\_\_

Language 2 \_\_\_\_\_

Language 4 \_\_\_\_\_

Language 5 \_\_\_\_\_

**Which language do you generally prefer to SPEAK in? Please list them in order of preference.**

Language 1 \_\_\_\_\_

Language 2 \_\_\_\_\_

Language 2 \_\_\_\_\_

Language 4 \_\_\_\_\_

Language 5 \_\_\_\_\_

**What cultures do you identify with? (E.g., Norwegian, British, American, etc.) Please list the cultures in the box below and then using the scale below from 0 to 10 please rate the extent to which you identify with each culture.**

List cultures here

Degree of identification

1 \_\_\_\_\_

\_\_\_\_ %

2 \_\_\_\_\_

\_\_\_\_ %

3 \_\_\_\_\_

\_\_\_\_ %

4 \_\_\_\_\_

\_\_\_\_ %

5 \_\_\_\_\_

\_\_\_\_ %

**Do you feel that you were once better in one of your languages and that you have become less fluent?**

(1)  Yes

(2)  No

**If yes, which one?**

\_\_\_\_\_

**At what age?**

—

**In which language do you usually do the following tasks?**

Count, add, multiply, and do simple mathematics \_\_\_\_\_

Dream \_\_\_\_\_

Express anger or affection \_\_\_\_\_

Talk to yourself \_\_\_\_\_

**When you are speaking, do you ever find yourself accidentally mixing words or sentences from the various languages you know?**

- (1)  Never
- (2)  1
- (3)  2
- (4)  3
- (5)  4
- (6)  Half of the time
- (7)  6
- (8)  7
- (9)  8
- (10)  9
- (11)  All the time

- (5)  4
- (6)  Half of the time
- (7)  6
- (8)  7
- (9)  8
- (10)  9
- (11)  All the time

**If yes, how often does English intrude into your Norwegian?**

- (1)  Never
- (2)  1
- (3)  2
- (4)  3

**And how often does Norwegian intrude into your English?**

- (1)  Never
- (2)  1
- (3)  2
- (4)  3
- (5)  4
- (6)  Half of the time
- (7)  6
- (8)  7
- (9)  8
- (10)  9
- (11)  All the time

**If there is anything else that you feel is interesting or important about your language background or language use, please comment below.**

\_\_\_\_\_

\_\_\_\_\_

**Section 2: All questions below refer to your knowledge of NORWEGIAN**



Which dialect of Norwegian do you speak?

---

Please list the number of years and months you spent in each environment below:

	Years	Months
In a country where Norwegian is spoken	—	—
In a family where Norwegian is spoken	—	—
In a school and/or work place where Norwegian is spoken	—	—

On a scale where 0 is "not a contributor", and 10 is "most important contributor", please indicate how much the following factors contributed to how you learned Norwegian.

Interacting with family	—
Interacting with friends	—
Reading (e.g., books, magazines, on-line)	—
School/education	—
Watching TV/streaming	—
Listening to music/media	—

Using a scale where 0 is "never" and 10 is "all the time", please rate to what extent you are currently exposed to Norwegian in the following contexts.

Interacting with family	—
Interacting with friends/colleagues	—
Reading (e.g., books, magazines, on-line)	—
Language courses/self-instruction	—

Watching TV/streaming	—
Listening to music/media	—

**Please indicate your level of proficiency in SPEAKING Norwegian.**

- (1)  None
- (2)  Very low
- (3)  Low
- (4)  Fair
- (5)  Slightly less than adequate
- (6)  Adequate
- (7)  Slightly more than adequate
- (8)  Good
- (9)  Very good
- (10)  Excellent
- (11)  Perfect

**Please indicate your level of proficiency in UNDERSTANDING Norwegian.**

- (1)  None
- (2)  Very low
- (3)  Low
- (4)  Fair
- (5)  Slightly less than adequate
- (6)  Adequate
- (7)  Slightly more than adequate
- (8)  Good

- (9)  Very good
- (10)  Excellent
- (11)  Perfect

**Please indicate your level of proficiency in  
READING Norwegian.**

- (1)  None
- (2)  Very low
- (3)  Low
- (4)  Fair
- (5)  Slightly less than adequate
- (6)  Adequate
- (7)  Slightly more than adequate
- (8)  Good
- (9)  Very good
- (10)  Excellent
- (11)  Perfect

**Please indicate your level of proficiency in  
WRITING Norwegian.**

- (1)  None
- (2)  Very low
- (3)  Low
- (4)  Fair
- (5)  Slightly less than adequate
- (6)  Adequate
- (7)  Slightly more than adequate
- (8)  Good
- (9)  Very good
- (10)  Excellent
- (11)  Perfect

**Please indicate your level of proficiency in  
Norwegian GRAMMAR.**

- (1)  None
- (2)  Very low
- (3)  Low
- (4)  Fair
- (5)  Slightly less than adequate
- (6)  Adequate
- (7)  Slightly more than adequate
- (8)  Good
- (9)  Very good

- (10)  Excellent
- (11)  Perfect

**Please indicate your level of proficiency in  
Norwegian VOCABULARY.**

- (1)  None
- (2)  Very low
- (3)  Low
- (4)  Fair
- (5)  Slightly less than adequate
- (6)  Adequate
- (7)  Slightly more than adequate
- (8)  Good
- (9)  Very good
- (10)  Excellent
- (11)  Perfect

**Please indicate your level of proficiency in  
Norwegian SPELLING.**

- (1)  None
- (2)  Very low
- (3)  Low
- (4)  Fair
- (5)  Slightly less than adequate
- (6)  Adequate
- (7)  Slightly more than adequate
- (8)  Good
- (9)  Very good
- (10)  Excellent
- (11)  Perfect

**On a scale of 1 to 10, how strongly regional is  
your spoken Norwegian in your opinion?**

—

**On a scale where 0 is "never" and 10 is "all the  
time", please rate how frequently others identify  
which part of Norway you come from when they  
hear you speaking.**

—

How important is speaking in your own dialect for you?

- (1)  Not at all
- (2)  1
- (3)  2
- (4)  3
- (5)  4
- (6)  Moderately important
- (7)  6
- (8)  7
- (9)  8
- (10)  9
- (11)  Extremely important

To what extent would you say you modify your own dialect when speaking to a person with a different dialect?

- (1)  Not at all
- (2)  1
- (3)  2
- (4)  3
- (5)  4
- (7)  Moderately
- (8)  6
- (9)  7
- (10)  8
- (11)  9
- (12)  Totally

**Section 3: All questions below refer to your knowledge of ENGLISH**

What kind of accent do you think your spoken English has? (e.g., British/American/other/none in particular)

---

To the best of your knowledge, please give the age when you .....

- Began acquiring English      —
- Became fluent in speaking English      —
- Started learning to read English      —
- Became fluent in reading English      —

Please list the number of years and months, if any, that you spent in each environment below:

Years

- In a country where English is spoken      —
- In a family where English is spoken      —
- In a school/work place where English is spoken      —

On a scale where 0 is "not a contributor", and 10 is "most important contributor", please indicate

how much the following factors contributed to how you learned English.

- Interacting with family \_\_\_\_\_
- Interacting with friends \_\_\_\_\_
- Reading (e.g., books, magazines, on-line) \_\_\_\_\_
- School/education \_\_\_\_\_
- Watching TV/streaming \_\_\_\_\_
- Listening to music/media \_\_\_\_\_

On a scale where 0 is "never" and 10 is "all the time", please rate to what extent you are currently exposed to English in the following contexts.

- Interacting with family —
- Interacting with friends —
- Reading (e.g., books, magazines, on-line) —
- School/education —
- Watching TV/streaming —
- Listening to music/media —

On the scale below, please indicate your level of proficiency in SPEAKING English.

- (1)  None
- (2)  Very low
- (3)  Low
- (4)  Fair
- (5)  Slightly less than adequate
- (6)  Adequate
- (7)  Slightly more than adequate
- (8)  Good
- (9)  Very good
- (10)  Excellent
- (11)  Perfect

On the scale below, please indicate your level of proficiency in UNDERSTANDING English.

- (1)  None
- (2)  Very low
- (3)  Low
- (4)  Fair
- (5)  Slightly less than adequate
- (6)  Adequate
- (7)  Slightly more than adequate
- (8)  Good
- (9)  Very good
- (10)  Excellent
- (11)  Perfect

On the scale below, please indicate your level of proficiency in READING English.

- (1)  None
- (2)  Very low
- (3)  Low
- (4)  Fair
- (5)  Slightly less than adequate
- (6)  Adequate
- (7)  Slightly more than adequate
- (8)  Good
- (9)  Very good
- (10)  Excellent
- (11)  Perfect

On the scale below, please indicate your level of proficiency in WRITING English.

- (1)  None
- (2)  Very low
- (3)  Low
- (4)  Fair
- (5)  Slightly less than adequate
- (6)  Adequate
- (7)  Slightly more than adequate
- (8)  Good
- (9)  Very good
- (10)  Excellent
- (11)  Perfect

**Please indicate your level of proficiency in English**

**GRAMMAR.**

- (1)  None
- (2)  Very low
- (3)  Low
- (4)  Fair
- (5)  Slightly less than adequate
- (6)  Adequate
- (7)  Slightly more than adequate
- (8)  Good
- (9)  Very good
- (10)  Excellent
- (11)  Perfect

**Please indicate your level of proficiency in English**

**VOCABULARY.**

- (1)  None
- (2)  Very low
- (3)  Low
- (4)  Fair
- (5)  Slightly less than adequate
- (6)  Adequate
- (7)  Slightly more than adequate
- (8)  Good
- (9)  Very good
- (10)  Excellent
- (11)  Perfect

**Please indicate your level of proficiency in English**

**SPELLING.**

- (1)  None
- (2)  Very low
- (3)  Low
- (4)  Fair
- (5)  Slightly less than adequate
- (6)  Adequate
- (7)  Slightly more than adequate
- (8)  Good
- (9)  Very good
- (10)  Excellent
- (11)  Perfect

**Please indicate your level of proficiency in English**

**PRONUNCIATION.**

- (1)  None
- (2)  Very low
- (3)  Low
- (4)  Fair
- (5)  Slightly less than adequate
- (6)  Adequate
- (7)  Slightly more than adequate
- (8)  Good
- (9)  Very good
- (10)  Excellent
- (11)  Perfect

**In your perception, how much of a Norwegian accent do you have when you speak English?**

- (1)  None at all
- (2)  1
- (3)  2
- (4)  3
- (5)  4
- (6)  Moderate accent
- (7)  6
- (8)  7
- (9)  8
- (10)  9
- (11)  Extremely strong accent

**Please rate how frequently others identify you as a non-native speaker because of your accent when speaking English.**

- (1)  Never
- (2)  1
- (3)  2
- (4)  3
- (5)  4
- (6)  Half of the time
- (7)  6
- (8)  7
- (9)  8
- (10)  9
- (11)  All the time

**How important it is for you to have a good accent when speaking English?**

- (1)  Not at all
- (2)  1
- (3)  2
- (4)  3
- (5)  4
- (6)  Moderately important
- (7)  6
- (8)  7
- (5)  4
- (6)  Moderate effort
- (7)  6
- (8)  7
- (9)  8
- (10)  9
- (11)  Constant effort

- (9)  8
- (10)  9
- (11)  Extremely important

**How much effort have you put into improving your accent?**

- (1)  No effort at all
- (2)  1
- (3)  2
- (4)  3

## Appendix B

### PARTICIPANT INFORMATION SHEET AND CONSENT FORM

#### INVITATION TO PARTICIPATE IN A RESEARCH PROJECT

#### THE PROCESSING OF ENGLISH AS A SECOND LANGUAGE

We are looking for Native speakers of Norwegian to take part in a language study investigating the processing of English as a second language.

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 and hearing and have no diagnosed language impairments such as dyslexia or stuttering.

The study has three components.

1. A language background questionnaire, which you will be asked to complete on-line and should take no more than 20 minutes.
2. A series of simple language tests which you will be invited to participate in following the completion of the questionnaire. These tests will take approximately 1 hour.
3. An eye-tracking study in which we will measure your eye-movements as you read simple English sentences. This experiment will take approximately 30 minutes.

The principal investigators managing this study at the Experimental Linguistic Laboratory (UiA) are:

Professor Linda Wheeldon ([linda.r.wheeldon@uia.no](mailto:linda.r.wheeldon@uia.no)) & Professor Allison Wetterlin ([Allison.wetterlin@uia.no](mailto:Allison.wetterlin@uia.no))

The studies will be run by 4 Master students at UiA, Kristiansand campus as part of their Master studies:

Bjørn Handeland ([bjørn.handeland@uia.no](mailto:bjørn.handeland@uia.no))

Beatrice Zitond Urland ([zitond16@uia.no](mailto:zitond16@uia.no))

Dag Haugland ([dag.haugland@uis.no](mailto:dag.haugland@uis.no))

Simon Wigstøl Olesen ([simwigol@gmail.com](mailto:simwigol@gmail.com))

Please contact us if you have any queries about the study.

#### WHAT IS THE STUDY ABOUT?

This study is designed to investigate aspects of the use of English as a second language, in particular, in the use of English by speakers that have Norwegian as their first language. We are interested in how aspects of a bilingual learning and language-use environment relate to different levels of second language English skills. The study has three components.

- An on-line 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 20 minutes to complete.
- A series of simple language tests that involve short, reading, vocabulary and spelling tests as well as some tasks involving repeating nonsense words. These tests will take approximately 1 hour to complete.
- An eye-tracking study that measures your eye-movements as you silently read simple English sentences. The eye-tracking study is done at the Experimental Linguistics Laboratory at UiA Kristiansand in an eye-tracking booth. During the eye-tracking study you will sit comfortably with your chin on a chin rest and silently read sentences displayed on a computer screen. Your eye-movements will be recorded automatically and at a distance by the eye tracker that uses a harmless beam of infrared light to track the movement of your eyes. Nothing will touch your eyes. This procedure is completely harmless. There are no risks or disadvantages associated with participation in the study.

If, after having read the information below, you decide to take part in the study please complete the consent form below and the questionnaire on this website. Once you have done this you will be contacted by e-mail by one of the researchers listed above. If you are found to be qualified as a participant, we will book times for you to complete the other sections of the experiment. Ideally you would finish all aspects of the study within a week or two of having completed the questionnaire.

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 and therefore you are essentially anonymous. 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 PC. Only the laboratory directors and experimenters will have access to the key relating your data number to your name.

#### VOLUNTARY PARTICIPATION AND THE POSSIBILITY 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. 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. If you at a later point, wish to withdraw consent or have questions regarding the project, you can contact the principle investigators managing this project (see e-mail addresses above).

#### 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. You have the right to access which information is recorded about you and the right to stipulate that any error in the information that is recorded is corrected.

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

The principal investigators have 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 will be anonymised or deleted a maximum of 5 years after the project has ended.

#### FINANCE

In appreciation for your time and effort, you will receive a voucher for Kafé Kampus for 200 NOK on completion of the three sections of this study. No payment will be received for partial participation.

#### APPROVAL

The Project is approved by the Regional Committee for Medical and Health Research Ethics [*insert reference number from REC (20xx/yyyy)*].



CONSENT FOR PARTICIPATING IN THE RESEARCH PROJECT

I AM WILLING TO PARTICIPATE IN THE RESEARCH PROJECT

**TITLE:** THE PROCESSING OF ENGLISH AS A SECOND LANGUAGE

- 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. 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)

## Appendix C

### ELL1 Language tests protocol

Participants name:

Participants number:

#### LEAP-Q

##### PREPARATION

- Upon entering the testing room, make sure all food and drinks are left outside.
- All cell phones must be turned off and put away. There is no signal inside the lab.
- Make sure all equipments are in their designated spots (computer, keyboards, mic)

##### 1. SENTENCE READING

- Check good quality recording – make sure the participant produces a fluent sentence. The sentence can be repeated if the participant did not get it right the first time.
- Participant may press - (minus key) for the sentence to reappear again.

*The winner of the race won a very large prize and the three losers cursed their bad luck.*

##### 2. YARK- Word Reading

- There are seven lists of words. The participant will read the instruction and notify the experimenter when he or she has finished. The experimenter will press the spacebar and begin the test.
- During the test:  

<b>Experimenter</b>	Button press to end recording
	Enter number of errors
	Button press to start next set of words

**List 1:** see, look, play, was, like, this, next, house, going, bell

**List 2:** hang, stand, their, living, again, first, slowly, score, found, bread

**List 3:** scream, journey, suppose, yawned, should, tissue, caught, stretching, tongue, copies

**List 4:** medicine, strengthen, source, creative, material, eventually, hygiene, despite, calm, journalism

**List 5:** excitable, dehydration, persuade, aggrieved, originate, courageous, atmospheric, familiarize, scenic, recurrence

**List 6:** ferocious, cynical, excursion, coincidental, abysmal, endeavor, rheumatism, haemorrhage, liaison, pseudonym

**List 7:** lacerate, bureaucracy, endogenous, coerce, archaic, facetious, pharmaceutical ochre, fruition, paediatrician

##### 3. SPELLING

- The participant will read the instruction and notify the experimenter when he or she has finished. The experimenter will press the spacebar and begin the test.
- There are 20 words. The words will be played one at a time. The participant can repeat the audio recording by pressing the enter key.
- During the test, the experimenter does not have to press any buttons.

#### 4. BPVS3- SPOKEN WORD COMPREHENSION

- The participant will read the instruction and notify the experimenter when he or she has finished. The experimenter will press the spacebar and begin the test.
- For each word, there will be four picture choices and the participant must choose the correct one.
- There are four pauses in between.

#### 5. SOUND DROPPING

- The participant will read the instruction and notify the experimenter when he or she has finished. The experimenter will press the spacebar and begin the test.
- After each word is played, the experimenter need to:
  1. Press 0 to stop
  2. Enter errors (0 for no errors and 1 for errors)
  3. Press 0 again for the next word to appear

li.ɔgs	→	i.ɔgs
θaʊk	→	aʊk
'zæblət	→	æblət
tweln	→	teln
'splɔrtəl	→	'slɔrtəl
'skreɪpʊs	→	'skeɪpʊs
'plɑ:təf	→	'plɑ:tə
ja'lu:m	→	ja'lu:
'træs.dʒɔɪb	→	'træ.dʒɔɪb
'æb.sʌmpt	→	'æb.sʌpt
klɔ:p	→	klɔ:p
dʒɪlk	→	dʒɪk
'fi:kne	→	'fi:kə
'bi:lɪrɒm	→	'bi:lɪrɒm
'læn.spɒŋ	→	'læn.pɒŋ
'pɪp.sɔɪ	→	'pɪp.sɔɪ
'rɛmp.sɪlf	→	'rɛp.sɪlf
'wɔɪft.nʌp	→	'wɔɪf.nʌp

#### 6. NONSENSE SOUND REPETITION

- The participant will read the instruction and notify the experimenter when he or she has finished. The experimenter will press the spacebar and begin the test.
- After each word is played, the experimenter need to:
  1. Press 0 to stop
  2. Enter errors (0 for no errors and 1 for errors)
  3. Press 0 again for the next word to appear

/tʃʌz/  
/vɜrd/  
/slɔɪp/

/'θrɑ:n.tɪb/  
 /'wɛŋ.klɒft/  
 /'dʒɪŋ.prɔɪz/  
 /sploɪ.'tɛl/  
 /kɪf.'geɪ.brʌm/  
 /'rɛ.frum.taɪ/  
 /wɪ.kə'vu:n/  
 /'taɪ.ðə.spə,zɪk/  
 /,kɔ.dʒɪ'ti:θəl/  
 /fɑɪ'wɔ:k.pə,tʃɛz/  
 /'ret.kɪp,təʊ.vum.təm/  
 /ʃə'ti:sɑ,bʌnəl/  
 /,mən.ləʊ'si:.oʊ.frʌpt /  
 /'dʒɑ:.bə.,rɛm.jə.plʊv.zə/  
 /dʒɑɪ.'kɛn.tju.əl,vɪk.ləm/  
 /'vʊk.təl,rɔɪ.wən,ɔɪ.sə.,sæb/  
 /fən.'stu:.sɪb,lʌk.wəl,li.tɪz/  
 /'pəʊ.fə'skri:p.sə,stri:k,mɪl.ti.əs/  
 /'sɪl.væg.'su:f.jə.,wæ.spə.nə.trɪ,tʌm /

## 7. SENTENCE JUDGEMENT

- The participant will read the instruction and notify the experimenter when he or she has finished. The experimenter will press the spacebar and begin the test.
- The sentences will appear one at a time. The participants needs to press 0 for correct and 1 for incorrect.
- Only the grammar matters, not the spelling.

## 8. GSRT- STORY UNDERSTANDING

- The participant will read the instruction and notify the experimenter when he or she has finished. The experimenter will press the spacebar and begin the test.
- There are 8 stories and the participant has 25 minutes to complete the entire test. There are 5 questions (four multiple choice) for each story. Accuracy is more important than speed.
- If the participant did not finish all questions upon the allotted time, the experimenter needs to press – (minus key) to end the test.