ABSTRACT

Title of Thesis:	REPRESENTATIVE DATA AND PSYCHOMETRIC PROPERTIES OF SHORT VERSION OF THE KOREAN-ENGLISH BILINGUAL APHASIA TEST
	Seongsil Lee, Master of Arts, 2018
Thesis Directed By:	Dr. Yasmeen Faroqi Shah, Department of Hearing and Speech Sciences

Previous studies investigating the psychometric properties of the Bilingual Aphasia Test (BAT) have found variable results. This study sought to investigate performance of high proficiency Korean-English (KE) bilinguals on Korean and English BAT and examine the equivalency of test difficulty across the two languages. A total of thirty KE bilinguals took the Korean-BAT, English-BAT, and Korean-English Translation Test (KETT). Their performance was evaluated and compared across two languages. Results showed that KE bilinguals performed above 80% on all subtests, however, they displayed different performance between Korean and English in three subtests. Item analyses found eighteen items with whose accuracy was below 80% and sixteen item pairs with unequal performance across the two languages. These results support the importance of testing psychometric properties of BAT and developing normative data for each language. Based on the representative data, recommendations for further modification of the BAT and a new ceiling criterion are proposed.

REPRESENTATIVE DATA AND PSYCHOMETRIC PROPERTIES OF SHORT VERSION OF THE KOREAN-ENGLISH BILINGUAL APHASIA TEST

by

Seongsil Lee

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Arts
2018

Advisory Committee:
Dr. Yasmeen Faroqi Shah, Chair
Dr. Nan Bernstein Ratner
Dr. Min Wang

Acknowledgements

First, I would like to express my sincere gratitude to my committee chair and advisor, Dr. Yasmeen Faroqi Shah for her continuous support and motivation throughout this project. Dr. Shah has always inspired me to go above and beyond what I thought I was capable of.

I would also like to thank my other committee members, Dr. Nan Bernstein Ratner and Dr. Min Wang, for their contributions and suggestions that helped improve this project.

Finally, I would like to thank my parents Bungrae Lee and Hyeonsug Lee who always told me to follow what my heart desires. This project would have been impossible without their love and support.

Table of Contents

List of Tables

Introduction

It is important to assess performance in both languages of bilinguals with aphasia in order to obtain an accurate and complete picture of their strengths and deficits. Unfortunately, there are very few psychometrically valid standardized tests for bilingual aphasia assessment (Ivanova & Hallowell, 2013). Paradis (1989) developed the Bilingual Aphasia Test (BAT) to equally assess both languages of bilinguals. Currently, the BAT is available in 71 languages including Korean. According to Paradis (1989), different versions of the BAT in various languages are not merely translated from each other, but rather constructed to reflect cultural and linguistic differences in each language. However, the BAT lacks normative data and studies investigating the psychometric properties of the BAT are limited. So far, no study has investigated the psychometric properties of the Korean-BAT (K-BAT) nor established representative samples. Furthermore, currently available Korean aphasia tests are limited in their ability to assess both languages of Korean-English (KE) bilinguals. The absence of a tool that can reliably assess both Korean and English for KE bilinguals poses a significant challenge in clinical practice as well as in bilingual aphasia research. It is particularly important for a bilingual test to be comparable in content and difficulty in each language to minimize the risk of incorrectly concluding that aphasia is more severe in one language versus another. The current study addressed these gaps in knowledge by testing if a representative sample of highly proficient KE neurologically healthy bilinguals achieved ceiling performance and examining the equivalence of test difficulties across Korean and English test items.

The demand for a language assessment tool for KE bilinguals is increasing as there are an estimated 1.8 million Koreans in the United States (U.S. Census, 2014) and over 1.1 million people in the United States use Korean at home (U.S. Census, 2011). In addition to that, KE

bilinguals are common in South Korea since English is included as a mandatory subject in every public elementary, middle, and high school's curriculum. Moreover, English is one of the immigrant languages spoken in South Korea and nearly 63,000 English speaking immigrants reside in South Korea (Gary & Fennig, 2017).

The following section will provide a brief background on bilingual language performance with a focus on language performance of KE bilinguals. Next, the review of assessment of bilingual aphasia will be provided. An overview of the BAT will also be provided with a focus on studies investigating the psychometric properties of the BAT. Lastly, a discussion of specific scoring considerations for KE bilinguals' language performance will be given.

Bilingual Performance on Language Measures

Bilinguals perform qualitatively and quantitatively differently from monolinguals on several language measures (Bialystok, Craik, & Luk, 2008; Filippi, Leech, Thomas, Green, & Dick, 2012; Ivanova & Costa, 2008; Portocarrero et al., 2007; Rosselli et al., 2002). Studies investigating bilingual performance on lexical retrieval tasks (i.e., naming tasks) point out that bilinguals show differences in speed and accuracy of the naming responses. The review of literature shows that, compared to monolinguals, bilinguals take longer (Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Ivanova & Costa, 2008) and produce fewer correct responses (Bialystok et al., 2008; Gollan, Fennema-Notestine, Montoya, & Jenigan, 2007; Kohnert, Hernandez, & Bates, 1998; Roberts, Garcia, Desrochers, & Hernandez, 2002). Sadat, Martin, Alario, and Costa (2011) used a phrase level production task and found that this bilingual difference is also evident in speech production beyond single words.

The different performance of bilinguals has also been observed in the verbal fluency task. During the verbal fluency tasks, bilinguals produced significantly fewer words than monolinguals (Gollan, Montoya, & Werner, 2002; Portocarrero, Burright, & Donovick, 2007; Rosselli et al., 2000; Rosselli et al., 2002). However, the type of verbal fluency task is found to influence results. Generally, bilingual performance is similar to monolinguals for phonemic fluency, but is worse for semantic fluency (Gollan et al., 2002; Portocarrero et al., 2007; Rosselli et al., 2000).

The findings on language comprehension have been mixed. The review of studies provides evidence of bilingual difference in comprehension tasks (Filippi et al., 2012; Mayo, Florentine, & Buus, 1997; Shi, 2010). Bialystok and Luk (2011) compared English receptive vocabulary scores of monolinguals and bilinguals and found that the mean standard scores of bilinguals were lower than monolinguals, especially among younger adults. However, in a semantic categorization task, bilinguals have shown similar performance to monolinguals (Gollan et al., 2005). Findings on sentence comprehension in the presence of background noise have shown both worse performance (Mayo et al., 1997; Shi, 2010) and improved performance (Filippi et al., 2012). Lastly, the study by Anderson, Vanderhoff, and Donovick (2013) compared performance between monolingual and bilingual college students on writing tasks. They found a bilingual effect in writing. Bilinguals wrote fewer words than monolinguals and received lower scores.

The review of the literature provides considerable evidence that bilinguals perform differently than monolinguals on language measures. Consequently, clinical assessment tools designed for and normed on monolinguals cannot be directly applied to bilinguals. Therefore, in order to reliably assess both languages of bilinguals, a language assessment tool that is

specifically designed to assess bilinguals' language performance should be used with a normative sample including bilinguals (Paradis, 2004; Paradis & Libben, 1987). Currently, four standardized tests have been developed to assess Korean speakers with aphasia: the Korean Boston Naming Test (K-BNT; Kim & Na, 1999), Korean Western Aphasia Battery (K-WAB; Kim & Na, 2001), Korean Test for the Differential Diagnosis of Aphasia (KTDDA; Park, 2005), and Daegu Aphasia Diagnostic Examination (DADE; Jung, 2006). Although these Korean aphasia tests have corresponding English tests, one should be careful when using the Korean and English test pairs (i.e., the K-WAB and WAB) to assess the linguistic abilities of KE bilinguals since bilinguals perform differently from monolinguals on language measures. For example, the normative samples of both the K-WAB and WAB did not include KE bilinguals. Hence, the comparison of language abilities and deficits between Korean and English will be invalid in this case. Kim and Na (1999) also realized the problem and claimed that the Korean aphasia tests should be used with discretion if the tests are to be used for a Korean-speaking population in an English-speaking region.

Not surprisingly, the bilingual language performance is influenced by their proficiency in each language (Gollan et al., 2007; Kohnert et al., 1998). Gollan et al. (2007) investigated the naming performance of 29 aging Spanish/English bilinguals on the Spanish and English Boston Naming Test (BNT). They found that bilinguals' naming performance reflects their language history, especially their experience with word forms. Similarly, Kohnert et al. (1998) that examined 100 young Spanish and English bilinguals' performance on the BNT in both Spanish and English found a correlation between language history information obtained from the initial questionnaire and the BNT scores in each language. These results suggest that bilinguals' language proficiency plays an important role on their language performance.

Language performance of Korean-English bilinguals. There is very limited research investigating the language processing and production of KE bilinguals. Four out of seven studies found in the literature examined the language performance of KE bilingual children (Kang, 2012; Kim, 2009; Lee, Kim, & Yim, 2013; Wang, Park, & Lee, 2006). Studies that investigated phonological awareness (PA) in KE bilingual children (Kang, 2012; Kim, 2009) found a bilingual advantage in PA in both languages (Kang, 2012) and a positive inter-lingual relationship between PA in Korean and English (Kim, 2009, Wang et al., 2006), while the interlingual correlation for literacy is mixed (Kim, 2009, Wang et al., 2006). Compared to monolingual Korean 3-5 year olds, KE bilinguals have a smaller vocabulary size in each language, but no difference in non-word repetition abilities (Lee et al., 2013).

Three out of seven studies examined the language performance of KE adult bilinguals (Hapsburg & Bahng, 2006; Suh, 2017; Yoo & Kaushanskaya, 2012). Yoo and Kaushanskaya (2012) investigated the performance of English monolinguals and KE bilinguals on phonological memory tasks in English with varying difficulty levels (i.e., word-span task, digit-span task, and non-word repetition task). They found that while monolinguals generally outperformed bilinguals, the differences in phonological memory performance of the two groups depended on the task types and difficulty. Hapsburg and Bahng (2006) investigated English monolinguals' and KE bilinguals' abilities to use contextual cues in a speech perception task in the presence of background noise. Bilingual groups were divided into a moderately proficient group and a non-proficient group based on their self-reported English proficiency scale. The study results indicated that the English monolingual group was more efficient in using contextual cues than the KE bilingual group. Moreover, the moderately proficient bilingual group was able to use contextual cues to a higher degree compared to the non-proficient bilingual group. Lastly, Suh

(2017) investigated the acquisition of the nominal domain of Korean in KE heritage speakers. Participants showed difficulties only when they had to choose correct comitative case markers, and this error pattern was correlated with the age of onset of English.

Overall, the limited research on language processing in KE bilinguals seems to present one important conclusion—namely, that there is a significant correlation between English and Korean phonological awareness in KE bilinguals, suggesting the possibility of inter-lingual influence in KE bilinguals' language performance. Additionally the bilingual effect on word retrieval is replicated in KE bilinguals (Yoo & Kaushanskaya, 2012). Although morphosyntactic differences have been documented in various bilingual groups (e.g., Krause, Bosch & Clahsen, 2015), we did not find any literature on morphosyntactic language use patterns in KE bilinguals. As will be discussed later, the Bilingual Aphasia Test has a large proportion of items that rely on morphosyntactic knowledge.

The Assessment of Bilingual Aphasia

Assessing the languages of bilinguals with aphasia is quantitatively and qualitatively different from assessing the language of monolinguals (Mindt et al. 2008). In bilinguals with aphasia, the degree to which one language is impaired may be different from the degree to which another language is impaired. Studies have shown that bilingual PWA present with varied language impairments and recovery patterns (Fabbro, 2001; Fabbro & Frau, 2001; Koumanidi Knoph, 2011; Paradis, 2001). Lorenzen and Murray (2008) summarized seven recovery patterns found in bilingual aphasia (see Table 1). Furthermore, given that PWA experience different degrees of impairment across language modalities, components of language in which a bilingual PWA shows impairment may vary with languages. Therefore, a thorough linguistic evaluation of

bilinguals with aphasia must include assessments of both languages using tasks and stimulus items matched in difficulty to allow direct comparison of performance in both languages (Fabbro, 2001; Lorenzen & Murray, 2003; Marrero, Golden, & Espe-Pfeifer, 2002; Mindt et al., 2008). Ideally, lexical items need to be matched across the two languages for psycholinguistic variables such as word frequency and phonological complexity. Sentences need to be matched for structural complexity across languages.

Table 1. Language recovery patterns of bilingual aphasia.

Recovery pattern	Language characteristics
Parallel recovery	Recovery of languages parallels the premorbid relative abilities. If one language were stronger premorbidly, it would return to being stronger.
Differential recovery	One language is recovered much better than the other compared to premorbid abilities.
Antagonistic recovery	One language is initially available, and as the other language recovers, the initially available language disappears.
Alternating antagonism	Repetition of the above pattern with languages alternating in availability. This may occur within cycles ranging from 24 hr to several months.
Blending recovery	Uncontrollable mixing of words and grammatical constructions of two or more languages even when attempting to speak in only one language. This should not be confused with the common bilingual practice of code switching.
Selective aphasia	Language loss only in one language with no measurable deficit in the other.
Successive recovery	The recovery of one language before the other(s).

Bilingual Aphasia Test

The Bilingual Aphasia Test (Paradis & Libben, 1987) is one of the few tests designed for assessment of languages in bi- and multi-linguals with aphasia. It is more widely used and cited than the Multilingual Aphasia Examination, which is available for Spanish only (Benton & Hamsher, 1994; Rey, Sivan, & Benton, 1991). The intended purpose of the BAT was to equally assess two languages of bilinguals with aphasia using a consistent measure across languages "so as to reliably and validly determine to what extent and in which aspects one language might be better preserved than another" (Paradis, 2011, p. 428). The BAT has been used to serve other purposes such as to study recovery pattern of bilinguals with aphasia, to examine treatment

outcomes, and differentiate between neurological impairments (Ivanova & Hallowell, 2009; Juncos-Rabadan, 1994; Koumanidi Knoph, 2011; Krishnan & Mathew, 2017; Manuel-Dupont, Ardila, Rosselli, & Puente, 1992; Muñoz & Marquardt, 2008; Roberts, 2008). Gomez-Ruiz and Aguilar-Alonso (2011) found that the Spanish version of the BAT is useful in differentiating between aphasia, Alzheimer's disease, mild cognitive impairment, and normal aging. Peristeri and Tsapkini (2011) found that the Greek version of the BAT was sensitive to differentially diagnose language deficits by especially detecting morphological and syntactic deficits in Greek speakers with aphasia.

The BAT consists of three parts. Part A contains 50 questions examining a patient's premorbid history of bilingualism. Part B assesses specific language structures (phonemic, phonological, morphological, syntactic, lexical, semantic) through comprehension, repetition, judgment, propositionizing, reading, and writing tasks (Paradis & Libben, 1987, p. 19). Part C assesses translation ability in given language pairs such as English-Korean (e.g., Korean English Translation Test [KETT]). The test items in Parts B and C were not intended to be direct translations but to be cross-linguistically equivalent across languages (Paradis & Libben, 1987). For instance, issues with maintaining cross-linguistic equivalence across various versions are evident in the tests assessing phonology. In the E-BAT, test items assessing verbal auditory comprehension of minimal pairs include the words "cat", "mat", "fat", and "bat". If these words are directly translated into Korean, the translated set of words loses its ability to assess the same linguistic capacity as they were intended to do. Therefore, instead of directly translating English words into their correspondents in Korean, four single-syllable words (e.g., /mul/ "water", /bul/ "fire", /sul/ "alcohol", and /k'ul/ "honey") were chosen to be presented as minimal pairs under the verbal auditory comprehension task in the K-BAT. Moreover, every language differs in its

morphosyntactic structures and the BAT was designed to accommodate these differences between languages so that the test could be matched in difficulty. For example, sets of standard sentences under the sentence comprehension subtest are modeled based on the sentence structure considered to be the simplest in a given language (i.e., SVO in English but SOV in Korean) (Paradis & Libben, 1987). For the nonstandard sentences, passive sentences were used for most languages, but for languages where passive sentences are too simple, artificial, or nonexistent, other structures that change the basic word order of the given language are used (Paradis & Libben, 1987).

The BAT is currently available in 71 languages, including Korean. These different language versions of the BAT were developed in collaboration with other researchers. Although the BAT was designed to assess a PWA's language exhaustively, but in a manageable manner, Paradis and Libben (1987) suggested that when it is not possible to administer the whole BAT due to time limitations or patient's impairment, the short version of the BAT can be administered instead. The short version consists of a select subset of 104 items from the original BAT. So far, two studies examined the performance of individuals with aphasia on the short version of the BAT (Ivanova & Hallowell, 2009; Krishnan & Mathew, 2017). The study by Ivanova and Hallowell (2009) found that the Russian short version of the BAT is effective in discriminating severity of impairment among Russian monolinguals with aphasia. The study by Krishnan and Mathew (2017) administered the Malayalam short version of the BAT on twenty-two Malayalam-English bilingual PWA and concluded that the Malayalam short version of the BAT has high test–retest reliability as well as content and construct validities. The present study used short versions of the E-BAT and K-BAT in examining their psychometric properties.

Psychometric Properties of the BAT

A clinically usable aphasia language test should be normed on a large sample of a neurologically healthy population (at least 100) with diverse age, education, and socio-economic status to obtain appropriate cut-off scores for normal performance (Franzen, 2003). Moreover, normative data based on the population with aphasia should also be obtained. The normative data is used to compute sensitivity and specificity of a test. Sensitivity refers to the percent of individuals with aphasia who perform below a cut-off score for normal performance (true positives) (Ross & Wertz, 2004). Specificity refers to the percent of individuals without aphasia who perform above the cut-off score for normal performance (true negatives) (Ross & Wertz, 2004). Both specificity and sensitivity examine the effectiveness of an aphasia test in discriminating PWA from people without aphasia. In reality, sensitivity and specificity often overlap to some degree. Thus, test developers should simultaneously consider both sensitivity and specificity of a test when deciding the ceiling criterion (Pepe, 2003; Strauss et al., 2006).

The BAT is limited in that the developers did not obtain normative samples and psychometric properties for the different languages. Instead, Paradis and Libben (1987) claimed that any person who has a practical command of two languages is expected to show ceiling performance (100%) in each subtest of the BAT. However, Paradis and Libben (1987) acknowledged that some subtests are harder than others and ceiling criterion of 100% would be too high for those subtests, especially for bilinguals who are not highly proficient in their languages. Thus, Paradis and Libben (1987) suggested a modified lowest possible ceiling approach, which allows ceiling criteria to be 80% for the following 19 subtests out of a total of 29 subtests: semicomplex and complex commands, verbal auditory discrimination, syntactic comprehension, semantic categories, synonyms, antonyms, grammaticality judgment, semantic

acceptability, sentence repetition, semantic opposite, derivational morphology, morphological opposites, mental arithmetic, listening comprehension, reading-sentences and text, dictation-sentences, and reading comprehension- words and sentences. Paradis and Libben (1987) argued that by applying the modified lowest ceiling approach, the BAT could identify language impairment in bilinguals without additional normative data for each language. However, studies investigating the psychometric properties of the BAT with neurologically healthy speakers found score discrepancies across languages for some subtests (Manuel-Dupont et al., 1992) or subtest performance that fell below the 80% modified ceiling (Muñoz & Marquardt, 2008). Both of these studies were conducted in Spanish-English bilingual speakers. Manuel-Dupont et al. (1992) examined the performance of 17 neurologically healthy Cuban-American adults on the English BAT, Spanish BAT, and English-Spanish translation test and found significant cross-linguistic differences on four subtests (sentence construction, number of words, morphological opposites, and reading). Similarly, Muñoz and Marquardt (2008) examined the performance of 22 neurologically healthy Spanish-English bilinguals on the short version of the English and Spanish BAT, and the English-Spanish translation test. Participants' scored higher on the English BAT than on Spanish BAT, which was consistent with their higher proficiency in English than Spanish. Additionally, item analysis revealed that the correct response rate for the 54 items was less than 70%, which falls well below the criterion suggested by Paradis and Libben (1987).

To summarize, very limited research has been conducted to obtain normative samples of the BAT and investigate the equivalence of test difficulty across languages. The results from two studies suggest that the modified ceiling approach alone is not enough in justifying the validity of the BAT by questioning its ability to measure language impairments in bilinguals with aphasia. Roberts (2008) also realized this gap in knowledge and argued that future research

should be conducted to obtain normative sample and test the equivalent difficulty of the BAT in different languages.

So far, no study has been conducted to obtain KE bilingual representative sample for the K-BAT, E-BAT and KETT to examine the modified lowest ceiling approach and determine appropriate ceiling criterion for normal performance and specificity value. Obtaining KE bilingual representative data and examining equivalence of test difficulties across two languages is essential to address the lack of an assessment tool that allows a direct comparison of the linguistic performance of KE bilinguals.

The current version of the K-BAT is not ideally constructed and could be improved in some ways. The wording of test instructions could be clearer, items could be organized by increasing difficulty, ambiguous picture stimuli could be replaced, and the font size of stimuli could be increased for readability. Moreover, while the full BAT is too long, the short BAT has too few test items. For these reasons, the K-BAT and E-BAT were first modified and then used for the present study. Modifications made to the K-BAT and E-BAT are explained in detail in Appendices I and II.

Scoring Considerations of the BAT for Korean-English Bilinguals

Like many other bilingual communities, code switching is common among both bilingual KE children and adults (Chung, 2006). Additionally, Korean and English differ in many linguistic aspects, which are briefly outlined below.

Phonological and phonetic characteristics. Unlike English, where stop consonants have two contrasting voicing features: voiceless and voiced, there is no voicing contrast in Korean. Instead, Korean stop consonants have three types of contrasts: aspirated, lax, and tense (Ha,

Johnson & Kuehn, 2009). All three stop contrasts are present at three distinct places of articulation: bilabial, alveolar, and velar. Korean does not have labiodental fricatives /f/, /v/ and linguadental fricatives /θ/, /ð/. Also, the /z/, /ʃ/, /ʒ/, /tʃ/, /dʒ/, and /r/ sounds do not exist in Korean (Ha et al., 2009). Kang and Guion (2006) found that the late KE bilinguals produce English voiceless stops as Korean aspirated stops and also produce English voiced stops similar to Korean lax and tense voiceless stops. The results suggest that KE bilinguals with varying degrees of proficiency may not differentiate five types of stop consonants in English and Korean as monolinguals. Korean has following ten vowels: /i/, /e/, /ɛ/, /y/, /ø/, /ɨ/, /ʌ/, /a/, /u/, and /o/, of which /y/, /ø/, and /ɨ/ vowels are not found in English (Ha et al., 2009). Moreover, unlike English vowels that is described as tense and lax (i.e., /i/-/ɪ/ and /ɛ/-/æ/), Korean do not have tense and lax distinctions and only have tense vowels (i.e., /i/ and /ɛ/). Studies have found that native Korean learners of English experienced difficulty in discriminating and producing /i/-/ɪ/ and /ɛ/-/æ/ vowel distinctions (Kim, 2010; Tsukada et al., 2005). The results imply that KE bilinguals, especially native Korean speakers who learn English as a second language, may show a lack of distinctions in certain vowel pairs.

Studies also show KE bilinguals differ from monolinguals in accentedness and prosody. Yeni-Komshian, Flege, and Liu (2000) investigated the pronunciation proficiency of KE bilinguals. Participants' age of arrival in the U.S. varied from 1 to 23 years. Heavier accents in English pronunciation were noted for the participants who arrived in the U.S. after the age of 5. The Korean pronunciation of participants who arrived in the U.S. before the age of 8 was observed to be heavily accented. Guion (2005) investigated the knowledge and implementation of English stress patterns by early and late KE bilinguals, and found that late KE bilinguals do not possess complete knowledge about English stress placement across lexical classes.

Sentence production. There are numerous morphosyntactic differences between Korean and English. Korean is Subject-Object-Verb (SOV) language, however, English is Subject-Verb-Object (SVO) language. Unlike English, Korean does not have auxiliary verbs and do not make subject-verb agreement. Lastly, prepositions and articles do not exist in Korean. Studies have found that Korea learners of English, even when they are advanced, have difficulty determining when to use the articles, 'the' and 'a' (Ko, Ionin, & Wexler, 2009; Park, 2005). Bitchener, Young and Cameron (2005) found that the use of English prepositions is especially challenging for English language learners whose L1 has no prepositions, such as Korean.

Two studies investigated the English oral and written narratives of Korean EFL learners (Kang, 2003, 2005). Kang (2003) analyzed oral "frog story" narratives of Korean adult EFL speakers, and found a heavy influence of Korean narrative strategies. Korean EFL learners produced considerably shorter narratives and included fewer explicit evaluative comments. Kang (2005) compared the English written narratives produced by Korean EFL learners and native English speakers. The study found that the English written narratives of Korean EFL learners were heavily influenced by Korean linguistic strategies and, therefore, deviated from the preferred written narrative style in the U.S. culture.

To summarize, the review of the literature provides significant considerations when assessing the language performance of KE bilinguals. Based on the distinctive Korean consonant and vowel systems, it was suggested that KE bilinguals may experience difficulties in pronouncing the five distinct types of English and Korean stop consonants (i.e., English voiced and voiceless stops, and Korean aspirated, tense, and lax stops) as well as certain English vowel pairs (i.e., /i/-/ɪ/, and /ɛ/- /æ/) (Ha et al., 2009; Kang & Guion, 2006; Kim, 2010; Tsukada et al., 2005). The English and Korean pronunciations of KE bilinguals may be accented depending on

their age of arrival (Yeni-Komshian et al., 2000). Even high proficiency KE bilinguals whose L1 is Korean may experience challenges in using English articles (Ko et al., 2009; Park, 2005) and prepositions correctly (Bitchener et al., 2005). Late KE bilinguals whose L1 is Korean may show erroneous English stress patterns (Guion, 2005). KE bilinguals often use code switching as a communicative strategy (Chung, 2006). Lastly, KE bilinguals' English oral and written narratives may be influenced by Korean linguistic strategies, therefore deviating from culturally appropriate English narrative styles (Kang, 2003, 2005). The presence of code switching behavior, inter-lingual influences in phonological awareness (Ha et al., 2009; Kang & Guion, 2006; Kim, 2010; Tsukada et al., 2005), and linguistic differences (Bitchener et al., 2005; Ko et al., 2009; Park, 2005) suggests the possibility that KE bilinguals will show distinctive error patterns in many language domains tested under the BAT. However, currently the BAT does not provide any guidelines on how to score KE bilinguals' performance by being sensitive to common errors patterns of KE bilinguals. Frequent code-switching behavior and accented pronunciation may affect their scores on the spontaneous speech subtest. Difficulties discriminating and producing distinct types of English and Korean consonants and vowels may influence their performance on the naming, verbal auditory comprehension, and translation of words subtests. Different word orders between English and Korean may affect their performance on the syntactic comprehension subtest. Lastly, difficulties in producing correct articles and prepositions may influence their scores on the translation of sentences and grammaticality judgment subtests.

Summary and Statement of the Problem

The review of literature outlines several key points related to the assessment of bilingual aphasia. First, bilingual performance differs from monolingual performance in lexical retrieval (Bialystok et al., 2008; Gollan et al., 2005, 2007; Ivanova & Costa, 2008; Kohnert et al., 1998; Roberts et al., 2002), verbal fluency (Gollan et al., 2002; Portocarrero et al., 2007; Rosselli et al., 2000, 2002), especially on semantic verbal fluency (Gollan et al., 2002; Portocarrero et al., 2007; Rosselli et al., 2000), sentence comprehension tasks (Mayo et al., 1997; Shi, 2010), writing (Anderson et al., 2013), and receptive vocabulary (Bialystok & Luk, 2011). The presence of phonological and morphosyntactic differences between Korean and English suggests the possibility that distinctive error patterns of KE bilinguals may influence their language performance (Bitchener et al., 2005; Ha et al., 2009; Kang & Guion, 2006; Kim, 2010; Ko et al., 2009; Park, 2005; Tsukada et al., 2005). Thus, monolingual aphasia tests and those norms cannot be directly applied to assess bilingual language performance. Moreover, the heterogeneous nature of language impairment and recovery patterns across languages (e.g. Fabbro, 2001) highlights the importance of matching test difficulties across different languages.

Although the BAT aims to address the need for a matched bilingual language assessment, studies investigating the psychometric properties of the BAT are limited and the results are preliminary (Ivanova & Hallowell, 2009; Juncos-Rabadan, 1994; Koumanidi Knoph, 2011; Krishnan & Mathew, 2017; Manuel-Dupont et al., 1992; Muñoz & Marquardt, 2008; Roberts, 2008). Contrary to Paradis and Libben's (1987) claim that neurotypical bilinguals can perform nearly at ceiling on the BAT with an applied modified lowest ceiling approach, Muñoz and Marquardt (2008) found less than 80% performance on many test items by neurotypical adults. In addition, studies examining the performance of neurologically healthy bilinguals on the BAT

found score discrepancies across languages (Manuel-Dupont et al., 1992; Muñoz & Marquardt, 2008).

Given the unresolved questions about the psychometric properties of the K-BAT and the equivalence of test difficulties between the K-BAT and E-BAT, the present study has three main goals. The first goal is to modify the K-BAT and E-BAT to improve its construction and to reflect recent developments in aphasiology. The second goal is to obtain KE bilingual representative data on the K-BAT, E-BAT, and KETT. The third goal is to examine the equivalence of test difficulties across two languages. Given that bilingual language performance is influenced by numerous variables, most notably age of L2 acquisition and language proficiency, this study seeks to minimize the influence of those variables by examining highly proficient sequential KE bilinguals.

Research Questions and Hypotheses

The present study examined the following questions regarding the performance of high proficiency KE bilinguals on the K-BAT, E-BAT and KETT:

1. How do high proficiency KE bilinguals perform on the modified E-BAT and K-BAT?

   It is predicted that high proficiency bilinguals will perform at ceiling on the BAT. Ceiling performance is defined as at least 80% performance on all subtests.

   If all high proficiency bilinguals perform at ceiling (i.e., 80%) on the BAT, then the 80% ceiling criterion would yield 100% specificity for the BAT.

   If the BAT was designed to test both languages equally, then performance of high proficiency KE bilinguals should not differ between the K-BAT and E-BAT for any of the subtests.

2. How do high proficiency KE bilinguals perform on the KETT?

   It is predicted that high proficiency bilinguals will perform at ceiling on the KETT, which is defined as 80% performance on all subtests.

   If all high proficiency bilinguals perform at ceiling (i.e., 80%) on the KETT, then the 80% ceiling criterion would yield 100% specificity for the KETT.

   If the KETT was designed to test the translation ability of both languages equally, then the performance of high proficiency KE bilinguals should not differ between Korean-to-English translation subtests and corresponding English-to-Korean translation subtests on the KETT.

3. Do the E-BAT, K-BAT, and KETT have balanced item difficulty?

   It is predicted that high proficiency KE bilinguals will perform near or at ceiling (e.g., 80%) on every item on the BAT.

   If the E-BAT and K-BAT have balanced item difficulty, then the correct response rate of high proficiency KE bilinguals on each item should not differ significantly from the correct response rate of its corresponding item across languages.

   If the KETT have balanced item difficulty, then the correct response rate of high proficiency KE bilinguals on each item on the Korean-to-English translation subtests should not differ significantly from the correct response rate of its corresponding item on the English-to-Korean translation subtests.

Methods

Experimental Design

A within-group design was used to address the research questions with high proficiency bilinguals. The outcome measures were participants' performance on the K-BAT, E-BAT, and KETT.

Participants

Thirty KE bilinguals (14 males, 16 females) were recruited from the University of Maryland College Park and its neighboring communities. Exclusion criteria for all participants were as follows: any history of speech and language impairments, neurological/cognitive deficits, or psychiatric conditions, less than high school education, proficient in languages other than Korean and English, and illiterate in Korean or English. All participants, except for one, were sequential bilinguals who learned Korean first. Only one participant was a simultaneous bilingual. Demographic information of the participants is reported in Table 2.

Table 2. Demographic characteristics of the participants

	Mean (SD)	Range
Age	25.3 (4.3)	18-34
Age of acquisition: English	8.6 (2.9)	0-13
Age of acquisition: Korean	0 (0)	0-0
Years of living in English spoken environment	7.9 (4.5)	3-18
Years of living in Korean spoken environment	15.5 (5.1)	4-20
Years of taking classes in English	9.6 (4.8)	3-20
Years of taking classes in Korean	11.7 (5.9)	0-20

Materials

Measures of language proficiency. Three measures were used to assess the language proficiency. English proficiency was measured using the Lexical Test for Advanced Learners of English (LexTALE; Lemhöfer & Broersma, 2012) which is a 5 minute vocabulary judgment test. In the present study, participants with LexTALE scores above 65% were considered as highly proficient English speakers. Korean proficiency was measured using a Korean lexical decision task that was developed with the same format at LexTALE. The Korean lexical decision task contained the same number of words and nonwords as the English LexTALE (e.g., 40 words and 20 nonwords). Korean words were matched to the word stimuli in the English LexTALE in word frequency, length (e.g., number of syllables in word) and/or class as closely as possible. Korean word frequency data are obtained from "the Frequency of Modern Korean Usage" published by the National Institute of Korean Language. The mean word frequency of Korean words (2497.7, SD = 2679.4) was not statistically different from the mean word frequency of English words (2521.1, SD = 2737.8) ($t(39) = 1.4$, $p = .14$). Similarly, the mean word length of Korean words (2.2, SD = .8) was not statistically different from the mean word length of English words (2.2, SD = .8) ($t(39) = .4$, $p = .66$). The list of stimuli for the Korean LexTALE are provided in Appendix IV. The task was presented to participants as a computer task using Microsoft PowerPoint. Each word was displayed on the screen for 2 seconds and participants were asked to verbally respond if the presented word is real Korean word or not. Lastly, the Bilingual Language Profile (BLP; Birdsong, Gertken, & Amengual, 2012) was used to obtain self-rated language dominance across the two languages. The BLP assesses different aspects of language dominance (e.g., language history, language use, language proficiency, and language attitudes) (Gertken, Amengual, & Birdsong, 2014) and yields a language dominance score of between -248

and 248. Language dominance score is calculated by subtracting one language total score from the other. Participants with dominance scores between -100 and +100 were considered as high proficient/balanced bilinguals. All participants obtained dominance scores between -100 and +100 from the BLP. Participants scored an average of 39.9 (SD=38.1, range = -27.06 to 97.2). The slight positive mean value indicated a small Korean dominance. Moreover, the mean Korean total score (167.9, SD = 19.9) was higher than the mean English total score (127.9, SD = 21.7) with statistical significant difference ($t(29) = -5.7$, $p < .01$). Similarly, the mean score was significantly higher in Korean LexTALE (96.9, SD = 3.2) than in English LexTALE (76.3, SD = 3.2) ($t(29) = 8.7$, $p < .01$). English LexTALE scores were not significantly correlated with years of living in an English spoken environment (Pearson r =.14) or years of taking classes in English (Pearson r = .20).

Bilingual Aphasia Test. The modified versions of the short E-BAT and K-BAT and the KETT were administered to each participant. Modifications were made to correct inaccuracies in instructions, replace ambiguous picture stimuli (N=10), and to increase item numbers (N=16) and difficulties. Modifications made to the K-BAT and E-BAT are explained in detail in Appendices I and II. For the purpose of the present study, the language background subtest of K-BAT/E-BAT was excluded, and the language history of each participant was obtained via the BLP because the BLP elicits a more detailed language history, has questions suitable for neurologically healthy bilinguals, and yields a dominance score. Excluding the language background subtest, the modified K-BAT and E-BAT included 95 items within the following 17 subtests: spontaneous speech, naming, pointing, simple and semi-complex commands, complex commands, verbal auditory discrimination, syntactic comprehension, repetition of words and nonsense words, repetition of sentences, series, verbal fluency, semantic opposites, reading,

copying, dictation, reading comprehension for words, and reading comprehension for sentences. The KETT included 58 items within the following 4 subtests: word recognition, translation of words, translation of sentences, and grammaticality judgment. The Korean to English translation subtests included 29 items. Similarly, the English to Korean translation subtests included 29 items.

Procedure

Participants were assessed at the University of Maryland College Park Aphasia Research Center or at home if participants could not be present at the testing site. All participants provided informed consent before the start of the study. All tasks were administered within one 2-hour session with a break in the middle. Participants were assessed first using three measures of language proficiency (e.g., the LexTALE, Korean LexTALE, and English-Korean BLP). Following this, the K-BAT, E-BAT, and KETT were administered. The order of K-BAT and E-BAT was counterbalanced across participants in order to minimize order effects, while KETT was always administered the last. Participants' responses were recorded using an audio recorder. The examiner, who is a graduate student in the Department of Hearing and Speech Sciences and is proficient in both English and Korean, administered all experimental tasks according to the administration instructions published in the K-BAT, E-BAT, and KETT. The same examiner scored participants' responses during testing. Additionally, using the audio-recorded data, the examiner double-scored their responses after the testing for the following subtests: spontaneous speech, naming, repetition of words and nonsense words, repetition of sentences, series, verbal fluency, semantic opposites, reading, word recognition, translation of words, translation of

sentences, and grammaticality judgment. Detailed scoring procedures of the E-BAT, K-BAT, and KETT are explained in Appendix III.

Data Analysis

To address the first research question, KE bilinguals' scores on the K-BAT and E-BAT were compared for each subtest using paired t-tests. A probability value of $p < .05$ was considered as statistically significant. In order to determine whether KE bilinguals performed near or at ceiling (i.e., 80%) on every subtest under the BAT, the mean subtest scores were compared with the cut-off scores of ceiling on each subtest. The second research question was addressed by comparing the performance of KE bilinguals' scores on the Korean-to-English direction subtests and English-to-Korean direction subtests of the KETT using paired t-test with a probability value of $p<.05$. In order to examine whether KE bilinguals performed near or at ceiling (i.e., 80%) on every subtest under the KETT, the mean scores of each subtest under the KETT were calculated and compared with the cut-off scores of ceiling. The third research question was addressed using differential item functioning (DIF) which derived an item difficulty (e.g., percent correct score) for each item on the K-BAT, E-BAT, and KETT (Swaminathan & Rogers, 1990). This was calculated as dividing the number of participants who produced a correct response over the total number of participants. The resulting item difficulty for each item was compared with the corresponding item across languages. Paired t-tests was used to investigate statistical differences between the two measures. A probability value of $p<.05$ was considered as statistically significant. Two subtests (i.e., spontaneous speech and verbal fluency) that contained items that were not dichotomously scored were excluded from the analysis since a percentage of correct response rate could not be calculated.

Results

In the following sections, the findings of E-BAT, K-BAT, and KETT will be presented by comparing the English and Korean performance. Next, the findings of item analysis will be presented.

Performance on the E-BAT and K-BAT

Mean scores and SD for each subtest on the K-BAT and E-BAT are reported in Table 3. Cut off scores indicating ceiling (80%) of each subtest are also reported in Table 3. All participants scored higher than ceiling (80%) for all subtests on E-BAT and K-BAT. The total score on BAT was higher in Korean than in English with statistically significant difference ($t(29) = -2.4$, $p = .02$) (see Table 3). This difference was mainly driven by the superior verbal fluency performance in Korean compared to English (see Table 3). Another significant score discrepancy was found in spontaneous speech subtest. English spontaneous speech performance was higher than Korean performance (see Table 3).

*Table 3. Subtest scores for the K-BAT and E-BAT, * = p<.05, ** = p <.01*

	English	Korean		Paired t-test	
	Mean (SD)	Mean (SD)	Cut off Score (80%)	t	p
Total Score	143.4 (6.0)	146.4 (6.4)	-	-2.4	0.02*
Spontaneous Speech	29.4 (.9)	28.9 (.4)	24	2.4	0.01*
Naming	9.7 (.4)	9.6 (.7)	8	0.6	0.54
Pointing	5.0 (.0)	5.0 (.0)	4	-	-
Simple Commands	6.0 (.0)	6.0 (.0)	4.8	-	-
Complex Commands	10.9 (1.3)	11.5 (.6)	9.6	-1.8	0.07
Verbal Auditory Discrimination	6.7 (.4)	6.8 (.3)	5.6	-0.7	0.48

Syntactic Comprehension	11.4 (.6)	11.6 (.5)	9.6	-1.2	0.20
Repetition Words	11.9 (.1)	11.9 (.1)	9.6	-	-
Repetition Sentences	2.9 (.2)	2.9 (.2)	2.4	-	-
Series	2.0 (.0)	2.0 (.0)	1.6	-	-
Verbal Fluency	20.5 (4.9)	23.4 (6.0)	-	-3.1	0.00**
Semantic Opposite	4.8 (.3)	4.7 (.4)	4	1.5	0.13
Reading	8.9 (.1)	9.0 (.0)	7.2	-1.0	0.32
Copying	2.0 (.0)	2.0 (.0)	1.6	-	-
Dictation	2.9 (.1)	3.0 (.0)	2.4	-1.0	0.32
Reading Comprehension Words	3.9 (.1)	3.9 (.1)	3.2	-	-
Reading Comprehension Sentences	3.9 (.2)	3.8 (.3)	3.2	1.0	0.32

Paired t-test was not conducted on the pairs with same mean scores across languages. This is indicated by (-).

As mentioned previously, the present study made modifications to the original short versions of E-BAT and K-BAT to increase item numbers and difficulties. A total of 12 items were added (6 items for each E-BAT and K-BAT) on the following subtests: naming, complex commands, and syntactic comprehension. Additionally, two items on the verbal auditory discrimination subtests were modified to increase difficulty (see Appendixes I & II). In order to examine the possibility that increased item numbers and difficulties could influence KE bilinguals' performance on the BAT, mean scores and SD for each modified subtest were recalculated after removing modified items. Resulting mean scores and SD for each modified subtest are reported in Table 5 along with the corresponding cut-off scores indicating ceiling (i.e., 80%). Since no modification was made to the KETT, the test was excluded from this analysis.

No significant change on performance was observed compared to the modified BAT performance described in Table 3. All participants scored higher than ceiling (80%) for all subtests even after removing modified items. No significant difference was observed between English and Korean performance on the four subtests (see Table 4).

Table 4. Subtest scores for the K-BAT and E-BAT after removing modified items

	English	Korean		Paired t-test	
	Mean (SD)	Mean (SD)	Cut off Score (80%)	T	p
Naming	5.9 (.1)	5.9 (.2)	4.8	1.0	0.32
Complex Commands	3.7 (.6)	3.9 (.3)	3.2	-1.6	0.11
Verbal Auditory Discrimination	5.7 (.4)	5.8 (.3)	4.8	-0.7	0.48
Syntactic Comprehension	9.5 (.6)	9.6 (.5)	8	-1.1	0.25

Performance on the KETT

Mean scores and SD for each subtest on the KETT are reported in Table 5 with corresponding cut off scores indicating ceiling (80%). Similar to the performance on the BAT, all participants scored higher than the expected ceiling (80%) for all subtests on the KETT. The total score on the KETT was higher in English to Korean direction than in Korean to English direction with statistically significant difference ($t(29) = -2.6$, $p = .01$) (see Table 5). This difference was mainly caused by the higher grammaticality judgment score in Korean compared to English (see Table 5).

*Table 5. Subtest scores for the KETT, * = p<.05*

	Korean to English	English to Korean		Paired t-test	
	Mean (SD)	Mean (SD)	Cut off Score (80%)	t	p
Total Score	43.5 (2.6)	45.0 (2.0)	39.2	-2.6	0.01*
Word Recognition	5.0 (.0)	5.0 (.0)	4	-	-
Translation of Words	8.7 (.6)	8.8 (.8)	8	-0.7	0.45
Translation of Sentences	16.6 (.9)	16.8 (.8)	14	-0.9	0.36
Grammaticality Judgement	13.1 (2.1)	14.3 (1.6)	12.8	-2.3	0.02*

Paired t-test was not conducted on the pairs with same scores across languages. This is indicated by (-).

Item Analysis

A total of 18 items out of a total of 244 items had an accuracy less than 80% on the K-BAT, E-BAT, and KETT (see Table 6). Of the 18 items, 1 was from the E-BAT, 3 from the K-BAT, and 14 from the KETT (7 items from English to Korean translation subtests, and 7 items from Korean to English translation subtests). Of particular interest are the 6 items that had a correct response rate of less than 60% (see Table 6).

Table 6. Test items with an accuracy below than ceiling (80%)

Item #	Subtest	Item	Correct response rate (%)

E-BAT			
63	Syntactic Comprehension	The truck is not pulled by the car.	76
K-BAT			
32 #	Naming	건전지 (battery)	76
63	Syntactic Comprehension	트럭이 자동차에 끌려가는 것이 아닙니다. (the truck is not pulled by the car)	76
86	Semantic Opposites	참된 (true)	73
KETT (E-K)			
448	Translation of Words	Razor (면도기)	73
456		Madness (광기)	43*
481	Translation of Sentences	그녀가 나에게 너의 사진을 보녀주었다. (she showed me your pictures)	78
482	Grammaticality Judgement	작년에 우리는 많은 발전을 만들었다. (N) (last year, we made many progress)	73
483		작년에 우리는 많은 발전을 만들었다, (N) (last year, we made many progress)	73
488		경찰이 나의 할머니를 위해 수색했다. (N) (policeman examined for my grandmother)	53*
489		경찰이 나의 할머니를 위해 수색했다. (N) (policeman examined for my grandmother)	53*
KETT (K-E)			
442	Translation of Words	여행가방 (Suitcase)	76
444		못생김 (Ugliness)	20*
498	Grammaticality Judgement	This month, John did many contracts. (N)	56*
499		This month, John did many contracts. (N)	30*
505		The thief asked my money. (N)	76
506		Ann will find to marry someone. (N)	73
507		Ann will find to marry someone. (N)	73

E-K= English to Korean translation subtests; K-E= Korean to English translation subtests. Test items with correct response rate less than 60% are indicated by (*). Item added during

modification is indicated by (#). Ungrammatical sentences on the grammaticality judgment subtests are indicated by (N).

E-BAT. For Item #63 "the truck is not pulled by the car" seven out of thirty participants pointed to the semantically reversed foil picture of the truck that was pulled by the car.

K-BAT. For item #32 (naming subtest) "건전지" /gənzənzi/ (battery) incorrect response of participants included "배터리" /bætəli/ (battery), which is a Korean loanword from English word "battery". This word was considered incorrect since another widely used word "건전지" /gənzənzi/ (battery) existed in Korean. For item #63 (syntactic comprehension subtest) "트럭이 자동차에 끌려가는 것이 아닙니다" (the truck is not pulled by the car) seven out of thirty participants pointed to the semantically reversed foil picture of the truck that was pulled by the car. Lastly, for item #86 (semantic opposite subtest) "참된" /tʃamdwæn/ (true) participants produced a wide variety of correct and incorrect responses. Correct responses included the following "거짓된" /gəzitdwɑn/ (false), "잘못된" /zalmotdwɑn/ (wrong), "그릇된" /gɯlətdwɑn/ (wrong), "그른" /gɯlɯn/ (wrong), and "헛된" /hətdwɑn/ (false). Incorrect responses included the following "못된" /motdwɑn/ (bad), "안된" /andwɑn/ (sorry), "시끄럽다" /sikɯlʌpda/ (noisy), "참되지 않은" (not true), and "나쁜" /nabɯn/ (bad).

KETT. For item #448 (translation of words subtest) "razor" seven out of thirty participants produced an incorrect translation "레이저" /leɪzə/ (laser) which is a Korean loanword from English word "laser". For item #456 "madness" seventeen out of thirty participants produced incorrect translation "화남" /hawnam/ (angry state). For item #442 "여행가방" /jəhaŋgabaŋ/ (suitcase) participants produced a wide variety of correct and incorrect responses. Correct responses included "travel bag", "luggage", and "baggage". Incorrect responses included "carrier", "travel backpack", "trunk", and "travel case". For item #444 "못생김" /motsæŋgim/

(ugliness) twenty-four out of thirty participants produced inaccurate translation “ugly”. For item #481 (translation of sentences subtest) “그녀가 나에게 너의 사진을 보여주었다” (she showed me your pictures) participants omitted one or two word groups. For both items #482-483 (grammaticality judgment subtest) “작년에 우리는 많은 발전을 만들었다” (last year, we made many progress) eight out of thirty participants incorrectly judged the stimulus as grammatically correct sentence. Similarly, for both items #488-489 (grammaticality judgment subtest) “경찰이 나의 할머니를 위해 수색했다 (policeman examined for my grandmother) fourteen out of thirty participants incorrectly judged the stimulus as grammatically correct sentence. Participants displayed particular difficulty on the grammaticality judgment subtest in English. For the following item pairs: #498-499 “this month, John did many contracts” and #504-505 “the thief asked my money”, participants showed an error pattern where they were unable to produce grammatically acceptable sentences after judging the stimuli as incorrect. For items #506-507 (grammaticality judgment subtest) “Ann will find to marry someone” eight out of thirty participants incorrectly judged the stimulus as grammatically correct.

The next step in item analysis was to compare the relative difficulty of corresponding items in E-BAT and K-BAT. Similarly in the KETT, the comparable English to Korean and Korean to English items were compared. These were done by means of paired t-tests comparing performance across corresponding items. For examples, the item #24 (naming subtest) “book” in E-BAT corresponds to the item #24 (naming subtest) “책” /tʃæk/ (book) in K-BAT. The item #458 (translation of sentences subtest) “박대통령은 영어로 아침에 연설했다” (president Park spoke with English in the morning) in Korean to English subtest in the KETT corresponds to the item #470 (translation of sentences subtest) “professor lectured in French in the afternoon” in English

to Korean subtest in the KETT. A total of 16 item pairs showed unequal difficulties across two languages (see Table 7). Majority of these item pairs (12 out of 16) were part of the KETT.

*Table 7. Comparison of corresponding English and Korean items, * = p<.05, ** = p <.01*

				Paired t-test	
Item Pair #	Subtest	English Mean (SD)	Korean Mean (SD)	t	p
E-BAT and K-BAT					
E32-K32	Naming	.9 (.1)	.7(.4)	2.2	.03*
E47-K47	Complex Commands	3.5 (.9)	3.9 (.3)	-2.1	.04*
E48-K48	Verbal Auditory Discrimination	.8 (.3)	1 (.0)	-2.1	.04*
E86-K86	Semantic Opposites	.9 (.1)	.7 (.4)	2.5	.01*
KETT					
438-448	Translation of Words	1 (.0)	.7 (.0)	3.2	.00**
442-452		.7 (.4)	.9 (.1)	-2.2	.03*
444-454		.2 (.4)	1 (.0)	-10.7	.00**
446-456		.9 (.2)	.4 (.5)	4.3	.00**
461-473	Translation of Sentences	2.4 (.6)	2.8 (.3)	-3.0	.00**
482-498	Grammaticality Judgement	.5 (.5)	.7 (.4)	2.4	.02*
483-499		.3 (.4)	.7 (.4)	4.7	.00**
485-501		.8 (.4)	1 (.0)	2.6	.01*
488-504		.8 (.4)	.5 (.5)	-2.1	.04*
490-506		.7 (.4)	1 (.0)	3.2	.00**
496-512		.8 (.4)	1 (.0)	2.6	.01*

497-513	.8 (.4)	1 (.0)	2.6	.01*

Discussion

The aim of the study was to identify any subtests and test items with less than ceiling (i.e., 80%) and to investigate equivalence of test difficulty across two languages by examining the performance of KE bilinguals on K-BAT, E-BAT, and KETT. There are three primary findings of the present study. First, the study found that KE bilinguals scored higher than ceiling (80%) for all subtests on E-BAT, K-BAT, and KETT. However, different performance between English and Korean was observed in three subtests. Lastly, the item analysis identified 18 items that did not meet the ceiling level criterion of 80% as well as 16 item pairs that showed unequal performance across two languages indicating unequal test difficulty. In the following paragraphs, the results of the subtest performance on the experimental tasks will be discussed. Next, the discussion of an appropriate ceiling criterion and a specificity of the BAT will be provided. Different performance observed across two languages will also be discussed. Next, the results of the item analysis will be discussed with a focus on the observed error patterns of KE bilinguals and recommendations for further modification of the BAT. The discussion of language proficiency and the BAT performance will also be provided. Finally, conclusions and limitations of the current study will be discussed with suggestions for future directions in this area of research.

Subtest Performance on E-BAT, K-BAT, and KETT

The authors of the BAT (Paradis & Libben, 1987) claimed that the BAT could successfully differentiate language impairments of bilinguals with aphasia from neurologically

healthy bilinguals by applying a modified lowest ceiling approach which expects that non-impaired bilinguals could perform above 80% on the BAT. Consistent with the Paradis and Libben's (1987) assertion, the group of KE bilinguals performed above 80% on all subtests on E-BAT, K-BAT, and KETT (see Table 3 & 5). Performance above ceiling criterion suggests that any KE bilinguals who perform below 80% on the BAT can be assumed to have a language impairment if they have a neurological diagnosis (e.g. aphasia), or have low proficiency if they are neurotypical. However, it is important to consider the possibility that the ceiling criterion of 80% may fail to detect mild aphasia. Although Paradis and Libben (1987) had decided to apply the ceiling criterion of 80%, they argued that as far as possible, the ceiling criterion of 100% should be maintained. However, the present study found that KE bilinguals performed 100% on only four subtests (e.g., pointing, simple commands, series, and copying) (see Table 3). The results suggest that both 80% and 100% ceiling criteria are inadequate for the BAT to successfully differentiate language impairments of bilinguals and its varying severity. Thus, a new ceiling criterion should be proposed based on the representative data obtained in the present study. To summarize, the first two hypotheses regarding ceiling performance of KE bilinguals on subtests were confirmed.

Ceiling Criterion and Specificity

The specificity of BAT for KE bilinguals was calculated by dividing a number of non-impaired KE bilinguals who performed above the ceiling criterion by a total number of non-impaired KE bilinguals. This is reported in Table 8. The ceiling criterion was raised in steps of 5%. Unfortunately, sensitivity could not be computed since the present study did not recruit bilinguals with aphasia. Consistent with the prediction, the ceiling criterion of 80% suggested by

the authors of the BAT (Paradis & Libben, 1987) yielded a specificity of 100% for the BAT, except for the Korean to English translation subtests on the KETT (see Table 8). However, it is important to mention that the ceiling criterion of 90% also produced a specificity of 100% for both K-BAT and E-BAT (see Table 8). The results suggest that the 80% ceiling is too low for normal performance, thus could fail to detect varying severity of aphasia, especially mild aphasia. As mentioned earlier, in reality, the performance of non-impaired individuals overlaps with the performance of individuals with disorder to a considerable degree and choosing a ceiling criterion results in errors such as a false positive (e.g., people without disorder who failed the test) and a false negative (e.g., people with disorder who passed the test). Consequently, if the ceiling criterion of 80% is applied to the BAT, the false positive becomes zero, but the false negative increases greatly. In other words, if the ceiling criterion of 80% is applied, then all individuals with mild aphasia who perform above 80% on the BAT (e.g., false negative) will be considered as normal, which in turns decreases the sensitivity of the test. Hence, by applying the ceiling criterion of 80%, the BAT loses its ability to detect mild aphasia and cannot serve as a sensitive aphasia assessment of bilinguals. Based on the representative data obtained in the present study, therefore, it is recommended that the ceiling criterion of the K-BAT and E-BAT to be raised up to 90%. By applying the new ceiling criterion of 90%, the sensitivity of the BAT will be increased while keeping the 100% specificity.

On the other hand, the ceiling criterion of 80% produced a specificity of 93% on the Korean to English subtests and 100% on the English to Korean subtests on the KETT (see Table 8). If the ceiling criterion to be raised above 80%, then the specificity values for the KETT are decreased considerably. Since increasing the ceiling criterion above 80% causes the specificity values to drop, thus makes the KETT not sensitive in catching non-impaired individuals, it is

recommended that the ceiling criterion of 80% to be maintained for the KETT. However, further research is needed to examine the sensitivity of the BAT and KETT and investigate the interplay between two values.

Table 8. Specificity of the BAT and the KETT with varying ceiling criterion

	Specificity (%)			
Ceiling Criterion (%)	E-BAT	K-BAT	KETT (K-E)	KETT (E-K)
80	100	100	93*	100*
85	100	100	76.6	93
90	100*	100*	50	63
95	96	96	10	26
100	6	0	0	0

E-K= English to Korean translation subtests; K-E= Korean to English translation subtests. Appropriate specificity for each test are indicated by (*)

Unequal Performance between Korean and English

For equivalency of test difficulty across languages, the present study did not entirely support the authors' claim that BAT is matched in difficulty across languages. The group of KE bilinguals evidenced different performance on three subtests (e.g., spontaneous speech, verbal fluency, and grammaticality judgment) (see Table 3 & 5). The higher score on the English spontaneous speech subtest compared to Korean can be explained by the frequent code-switching behavior observed during Korean speech task, which decreased participants' scores. Moreover, the higher scores on the Korean verbal fluency and grammaticality judgment subtests compared to English likely reflect participants higher proficiency in Korean. In order to examine correlation between participants' BAT performance and factors contributing to language proficiency (i.e., age of acquisition, years of taking classes in Korean or English, years of living

in Korean or English speaking environment), Pearson correlations were performed. The additional analysis revealed that participants' performance on English verbal fluency subtest did not correlate with any factors contributing to language proficiency. On the other hand, performance on Korean verbal fluency subtest showed significant correlation with years of taking classes in Korean (Pearson $r = .4$, $p < .05$). Moreover, performance on English grammaticality judgment subtest showed significant correlation with years of taking classes in English (Pearson $r = .44$, $p < .05$). Similarly, performance on Korean grammaticality judgment subtest showed significant correlations with years of living in Korean speaking environment (Pearson $r = .54$, $p < .01$) and years of taking classes in Korean (Pearson $r = .45$, $p < .05$). The reason why performance on verbal fluency showed less correlation with factors influencing language proficiency is because verbal fluency task is known to be a measure of both language and executive function. Overall, the findings suggest the participants' performance on BAT subtests could have been influenced by individual factors contributing to language proficiency. More detailed discussion about language proficiency and BAT performance will follow.

The present study also identified 16 item pairs with unequal difficulties across two languages (see Table 7). The presence of cross-linguistic differences suggests that the BAT is not entirely matched in difficulty across languages, therefore equivalence of test difficulties across languages needs to be examined with non-impaired bilinguals' performance on the BAT. To summarize, three hypotheses regarding equivalence of test difficulty between two languages were rejected.

Item Analysis

In contrast to the findings of the subtest performance on the experimental tasks, item analysis found 18 items with an accuracy below the ceiling of 80% (see Table 6). The finding suggest that psychometric properties of BAT and its item difficulty require rigorous testing and verification with neurologically healthy speakers' performance for each language. The finding of the present study is consistent with prior finding of problematic items with less than 70% performance accuracy in neurologically healthy bilinguals (Muñoz & Marquardt, 2008). The present finding supports Roberts's (2008) assertion that research should be conducted to develop normative sample of BAT. To summarize, first hypothesis of third research question regarding item difficulty was rejected.

The 18 items with an accuracy below the ceiling (80%) can further be divided into three categories based on their error types: phoneme perception, dialectal difference, and item difficulty (see Table 9). The phoneme perception category includes problematic item for which error was caused by the failure to perceive subtle phonemic differences between English and Korean. The dialectal difference category includes items for which KE bilinguals produced distinct error patterns caused by dialectal differences. Lastly, the item difficulty category includes all the items that low accuracy was caused by inherent difficulties of the items. For the following paragraphs, each error pattern and related items will be discussed with recommendations for further modifications. At the end, additional recommendations for revising scoring criteria of spontaneous speech will also be discussed.

Phoneme perception. For item #448 "razor" on the translation of words subtest, KE bilinguals experienced difficulty perceiving phonemic distinction between two phonemes /r/ and /l/. Participants incorrectly translated "razor" into Korean loanword "레이저" /leɪzə/ (laser) from

English word “laser”. Since Korean consonant system does not have phonemic distinction between the phoneme /r/ and /l/, and only have phoneme /l/, participants who could not perceive phonemic difference between /r/ and /l/ in English considered the word “razor” as “laser” and produced corresponding Korean loanword “레이저” /leɪzə/ (laser) instead. Hence, the item #448 “razor” on the KETT is recommended to be replaced with another word that has same word class, frequency, and word length and does not require KE bilinguals to discriminate /r/ and /l/.

Dialectal difference. For item #32 “건전지” /gənzənzi/ (battery) on the naming subtest, KE bilinguals produced the word “배터리” /bætəli/ (battery) that is a Korean loanword from English word “battery” with similar pronunciation. Although the K-BAT did not provide “배터리” /bætəli/ (battery) as correct answer, it should be noted that both “건전지” /gənzənzi/ (battery) and “배터리” /bætəli/ (battery) are widely used to indicate battery among KE bilinguals. For item #442 “여행가방” /jəhaŋgabaŋ/ (suitcase) on the translation of words subtest, participants produced following incorrect translations: “캐리어” /kæliʌ/ (carrier) and “트렁크” /tɯləŋkɯ/ (trunk). These words exist in both English and Korean; however, have different meanings. Although the words “carrier” and “trunk” cannot indicate ‘suitcase’ in the U.S, both words are widely used to mean suitcase in Korea. Hence, it is recommended that the scoring criteria of both items should be expanded to account dialectal differences and to increase accuracy rate. For item #32 (translation of words subtest) “건전지” /gənzənzi/ (battery) on the K-BAT “배터리” /bætəli/ (battery) should also be included as correct answer. For item #442 (translation of words subtest) “여행가방” (Suitcase) both words “캐리어” /kæliʌ/ (carrier) and “트렁크” /tɯləŋkɯ/ (trunk) should be included as acceptable answers.

Item difficulty. Fifteen items are found to have accuracy rate less than ceiling (i.e., 80%) due to their inherent item difficulties (see Table 9). Among these fifteen items, eleven items are

recommended to be removed, one item to be replaced, and three items to be revised to decrease difficulty. First of all, both item #456 “madness” and item #444 “못생김” (Ugliness) on the translation of words subtests are recommended to be removed due to their extremely low accuracy rate (below 50%). Additionally, item #63 “the truck is not pulled by the car” on both E-BAT and K-BAT should be removed. The syntactic comprehension subtest included two negation sentences. One is the negation of the standard sentence (e.g., active sentence) “the truck does not pull the car” and the other is the negation of the type 1 non-standard sentence, which is the passive sentence in both English and Korean, “the truck is not pulled by the car”. Since the item #63, the negation of the non-standard passive sentences, had an accuracy below the ceiling (80%) in both languages, it could be considered as too difficult even for non-impaired KE bilinguals and not desirable to be included under the BAT. Lastly, items #482, 483, 488 and 489 as well as their corresponding items #498, 499, 504 and 505 on the grammaticality judgment subtests should be removed due to their inherent difficulty.

One item is recommended to be replaced. Item #86 “참된” /tʃamdwæn/ (true) on the K-BAT should be replaced with another Korean words with same meaning such as “진실된” /zinsildwæn/ (true) since majority of participants could not recognize the word “참된” /tʃamdwæn/ (true) and had trouble producing the semantic opposite.

Lastly, three items under the KETT are recommended to be revised to decrease difficulty. Item #481 (translation of sentences subtest) “그녀가 나에게 너의 사진을 보여주었다” (she showed me your pictures) should be modified to have less complex syntactic structure since participants experienced difficulty translating all the word groups. Hence, the sentence is recommended to be simplified to “그녀가 나에게 사진을 보여주었다” (she showed me pictures) or “그녀가 나의 사진을 보여주었다” (she showed my pictures). Additionally, the item #506-507 (grammaticality

judgment subtest) “Ann will find to marry someone” should also be modified to decrease its difficulty and to match difficulty with corresponding items on the opposite direction of translation (e.g., Korean to English). Therefore, the sentence is recommended to be revised to “Ann will find some to marry” similar to the corresponding item #490-491 “선희는 무엇인가 먹을 가지고 있다” (Sunhee has some to eat).

Table 9. Description of problematic items with recommendations

Item #	Phoneme Perception	Dialectal Difference	Item Difficulty	Recommendations
E-BAT				
63			•	Remove
K-BAT				
32		•		Expand scoring criteria
63			•	Remove
86			•	Replace
KETT (E-K)				
448	•			Replace
456			•	Remove
481			•	Revise
482			•	Remove
483			•	Remove
488			•	Remove
489			•	Remove
KETT (K-E)				
442		•		Expand scoring criteria
444			•	Remove
498			•	Remove
499			•	Remove
505			•	Remove
506			•	Revise
507			•	Revise

E-K= English to Korean translation subtests; K-E= Korean to English translation subtests.

Scoring criteria on spontaneous speech subtest. As mentioned previously, participants' lower scores on the Korean spontaneous speech subtest were greatly sacrificed due to their frequent code-switching behavior. However, code switching is common among KE bilinguals and often used as communicative strategy to facilitate conversation (Chung, 2006). Thus, scoring criteria on the spontaneous speech subtest need to be revised to reflect the pervasive code-switching phenomenon among KE bilingual. It is recommended that scoring criteria of code-switching (e.g., item #23) on both E-BAT and K-BAT to be revised from "very frequent (1 point) – absent (5 points)" to "very frequent (1 point) – occasional (5 points)". Additionally, participants' scores on the item #20 "articulation" were greatly decreased due to their accented pronunciation. It should also be noted that the scoring of accented pronunciation could have been biased by the examiner who is native speaker of Korean. The accented pronunciation of KE bilinguals is also common phenomenon and depends on their age of arrival (Yeni-Komshian et al., 2000). Thus, during the spontaneous speech task, speakers' intelligibility should be assessed instead of their articulation in order to reduce examiner's judgment bias and to account accented articulation as natural phenomenon among KE bilinguals.

Language Proficiency and BAT Performance

As previously mentioned, all participants, except for one, were sequential bilinguals who learned Korean first. The group of KE bilinguals showed Korean dominance along with their higher average score in Korean LexTALE than in English. Therefore, in order to examine the possibility that participants' language proficiency played a role on accuracy of individual items, correct response rate of each problematic items (N=8) on the E-BAT and the Korean to English translation subtests were recalculated excluding performance of nine participants who did not

reach the English LexTALE criterion satisfying high proficiency (65%). Nine participants (5 males, 4 females) with lower English proficiency scored an average of 59.7 (SD=2.23) on the English LexTALE. This additional analysis revealed mixed results. The accuracy of only one item increased to above ceiling (i.e., 80%) when speakers with lower English proficiency were excluded (see Table 10). Although the accuracy rates for two items (e.g., #506 and #507) were not increased above the ceiling of 80%, it was observed that the accuracy rates for these two items increased slightly. On the other hand, the accuracy rate of four items were decreased (see Table 10). Consequently, the results of this additional analysis could not confirm the effect of bilinguals' language proficiency on the BAT performance.

Table 10. Accuracy of English items before and after removing low proficiency data

Item #	Accuracy (%)	Accuracy (%) excluding low proficiency data
E-BAT		
63	76.6	76.1*
KETT (K-E)		
442	76	80**
444	20	19*
498	56	52*
499	30	23*
505	76	76
506	73	76**
507	73	76**

K-E= Korean to English translation subtests. Test items with decreased accuracy rate are indicated by (*). Test items with increased accuracy rate are indicated by (**).

In order to further investigate the effect of language proficiency on BAT performance, Pearson correlation analyses have been conducted between participants' English and Korean LexTALE scores and their respective BAT and KETT performance. The results indicated

significant correlation only between Korean LexTALE scores and performance on English to Korean subtests on the KETT (Pearson r =.36, $p < .05$). Korean LexTALE was not significantly correlated with K-BAT performance. Similarly, English LexTALE was not significantly correlated with E-BAT performance and Korean to English subtests performance on KETT. Additionally, linear regressions have been performed to further investigate what factors related to language proficiency (i.e., age of acquisition, years of taking classes in English or Korean, and years of living in English or Korean speaking environment) predict BAT performance. The linear regression analysis with K-BAT performance as dependent variable and years of taking classes in Korean and years of living in Korean speaking environment as factors found significant model ($F(3, 26) = 3.190$, $p < .05$, $R2 = .26$). However, none of the individual factors reached significance. Linear regression analyses with E-BAT and KETT in both directions as dependent variables found no significant results.

Overall, the findings suggest that the BAT performance is not significantly influenced by language proficiency of bilinguals, therefore, the BAT is not recommended to be used as a tool to assess language proficiency of bilinguals. However, given that the current study aimed to recruit high proficiency bilinguals, there was not a wide distribution of proficiency. Hence it is possible that testing a large number of participants with a wider range of proficiencies would reveal more consistent results showing relationship between language proficiency and bilinguals' BAT performance. Therefore, future research is warranted to further investigate the effect of language proficiency on the BAT performance by recruiting bilinguals with varying proficiency.

Conclusions and Future Directions

Three main conclusions can be drawn from the present study. First, KE bilinguals' performance above ceiling criterion of 80% on all subtests suggest that the modified lowest ceiling (80%) could potentially differentiate KE bilinguals with aphasia from individuals without aphasia. However, based on the representative data, the new ceiling criterion of 90% was suggested for both K-BAT and E-BAT in order to enhance the sensitivity of the tests in detecting varying severity of aphasia. The ceiling criterion of 80% was maintained for the KETT. Second, the presence of cross-linguistic differences suggests that the BAT is not entirely matched in difficulty across languages. Third, the item analysis identified 18 items with an accuracy below the ceiling of 80%. The results of the present study as well as findings of previous studies suggest that psychometric properties of the BAT and the equivalence of item difficulty across languages require testing with neurologically healthy speakers' performance. Moreover, based on the error types of items with low accuracy (N=18), 11 items are recommended to be removed, 2 items to be replaced, and 3 items to be revised to decrease difficulty, and 2 items to expand scoring criteria. Lastly, two scoring criteria of the spontaneous speech subtest are recommended to be modified to reflect common speech phenomena among KE bilinguals.

There were several limitations to the present study. It was suggested that a clinically usable aphasia language test should be normed on a large sample of a neurologically healthy population (at least 100) with diverse age, education, and socio-economic status to obtain appropriate cut-off scores for normal performance (Franzen, 2003). However, given the time constraint of this study, the sample size is relatively small. Based on the guidance, further examination of performance of neurologically healthy KE bilinguals on K-BAT, E-BAT, and KETT should continue to develop normative data. The present study was also limited in that the

study could not determine sensitivity as well as other psychometric properties of the BAT. In an effort to address this issue, further research should continue to investigate performance of KE bilinguals with aphasia with varying severity. The present study was also limited in that the study could not demonstrate the effect of bilinguals' language proficiency on the BAT performance. Given that the current study recruited a relatively small number of participants with a small distribution of proficiency, it is possible that a large number of participants with a wider range of proficiencies would have revealed different results. Therefore, future research should continue to further investigate the effect of bilinguals' language proficiency on the BAT performance by recruiting a large number of KE bilinguals with varying language proficiencies. Moreover, it is recommended that any future studies to use a different measure of language proficiency, other than LexTALE to thoroughly examine language proficiency and to further ascertain participants' pre-existing differences in language skills. Lemhöfer and Broersma (2012) also acknowledged the drawback and stated that LexTALE, which is a vocabulary knowledge test, could be used as a rough indication of proficiency when no other, more accurate measure is available.

Appendixes

Appendix I. Modification of English-Bilingual Aphasia Test.

Spontaneous Speech

Instruction for test administer was erroneous by including the wrong number of test items for the subtest. Therefore, the instruction was changed from “for the following five items” to “for the following six items”.

Naming Subtest

The number of test items for the naming subtest was increased from 6 to 10 by adding four additional items from original E-BAT. This modification was made since the subtest included limited number of items and especially lacks difficult items necessary to discriminate low proficiency bilinguals as well as mild to moderate PWA. Therefore, four difficult items that have longer syllable length (e.g., cigarette and thermometer) or low frequency (e.g., scissor and toothbrush) were added to increase the number of test items and difficulty so that the subtest could be sensitive to low proficiency KE bilinguals as well as to mild to moderate PWA. Four added items are followings: scissors, thermometer, cigarette, toothbrush.

Complex Command Subtest

The number of test items for the complex command subtests was increased from 1 to 3 by adding two additional items from original E-BAT. The reasons for this modification are same as those for naming subtest. Two added items are followings: “Here are three pencils, drop the yellow one on the floor, give me the blue one and pick up the red one” and “Here are three books, open the first one, turn over the second and pick up the third one”.

Verbal Auditory Discrimination Subtest

#45: Four picture drawings were replaced with picture drawings of ‘skate’, ‘bait’, ‘gate’ and ‘state’. This modification was made since original items of four picture drawings were not phonologically related to the target word ‘plate’. Therefore, in order to increase the difficulty, four picture drawings of phonological related distractors replaced the original items.

#46: Due to the ambiguity of the picture drawing of the word ’hip’, an arrow indicating the target word was inserted.

#47: Due to the ambiguity of the picture drawing of the word ‘grain’, it was replaced with the picture drawing of ‘brain’.

Syntactic Comprehension Subtest

Stimulus book page numbers did not match with the sentence stimuli. Therefore, the page number on the E-BAT were corrected to match the page number on the stimulus book.

#53: In order to increase the difficulty of the subtest, two additional items from the original E-BAT were added. Modification of stimulus book was not necessary since picture drawing of correct answers of the two added items were already included in the original stimulus page. Two added items are followings: ‘The cat is bitten by the dog’ and ‘It is the dog that the cat bites’.

Series Subtest

#75 was replaced with #261 “could you count from one to twenty-five?” from the original E-BAT in order to increase the difficulty of the test item.

Reading Subtest

The font size of the items in page 31 of the stimulus book was increased to be matched the font size of the same items under the K-BAT.

Reading Comprehension for Words

The font size of the target items under the subtest was increased to be matched the font size of the same items under the K-BAT.

#98: Due to the ambiguity of the picture drawing of the distractor #1, it was replaced with the picture drawing of a new distractor ‘lamp’. In addition, the picture drawing of the distractor #4 was replaced with the drawing of the new distractor ‘camp’ for the same reason.

#99: Due to the ambiguity of the picture drawing of the distractor #3, it was replaced with the picture drawing of a new distractor ‘hair’.

The following subtests under the original short E-BAT are not modified and used entirely for the modified E-BAT: pointing, simplex and semi-complex commands, repetition of words and nonsense words, repetition of sentences, verbal fluency, semantic opposites, copying, dictation, and reading comprehension for sentences.

Appendix II. Modification of Korean-Bilingual Aphasia Test.

Spontaneous Speech

Instruction for test administer was erroneous by including the wrong number of test items for the subtest. Therefore, the instruction was changed from "아래 제시된 다섯 개의 질문항목에 대하여" to "아래 제시된 여섯 개의 질문항목에 대하여".

Naming Subtest

The number of test items for the naming subtest was increased from 6 to 10 by adding four additional items from original K-BAT. This modification was made since the subtest included limited number of items and especially lacks difficult items necessary to discriminate low proficiency bilinguals as well as mild to moderate PWA. Therefore, four difficult items that have longer syllable length (e.g., 온도계 and 건전지) or low frequency (e.g., 가위 and 칫솔) were added to increase the number of test items and difficulty so that the subtest could be sensitive to low proficiency KE bilinguals as well as to mild to moderate PWA. Four added items are followings: 가위, 온도계, 건전지, and, 칫솔. The last item was corrected to its standard Korean from the item #287 '치솔' included in the original K-BAT.

Complex Command Subtest

The number of test items for the complex command subtests was increased from 1 to 3 by adding two additional items from original E-BAT. The reasons for this modification are same as those for naming subtest. Two added items are followings: "여기 연필 세 자루가 있습니다. 노란 연필은 바닥에 던지시고, 파란 연필은 저에게 주시고, 빨간 연필은 손으로 집어 주세요" and "여기 책이 세 권 있습니다. 첫번째 책은 열어 주시고, 두번째 책은 뒤집어 주시고, 세번째 책은 들어주세요". The last item was not translated fully in the original K-BAT. In other words, the item #47 included only the

first sentence of the complex command “here are three books”. Therefore, the equivalent complex command in the E-BAT was translated to Korean to create the complete item.

Verbal Auditory Discrimination Subtest

Instruction did not match with the items in the stimulus book. The instruction says that “if I say glass, you would touch picture #4 because it represent the glass”. However, the corresponding picture drawing of distractor #4 did not represent the glass. Therefore, correct instruction corresponding to the picture stimuli was borrowed from the original K-BAT. The replace instruction is “이제 제가 단어 하나를 소리내어 말하겠습니다. 이 단어에 해당되는 그림을 손가락으로 가리켜 주세요. 제가 말한 단어의 그림을 찾으실 수 없는 경우도 있습니다. 그런 경우에는 여기 있는 X자를 가리켜 주세요. 예를 들어, 제가 ‘머리’라고 말하면 3번 그림을 가리키시면 됩니다. 만약 ‘새’라고 말하면 ‘새’의 그림은 이 페이지에 없으므로 X자를 가리키시면 됩니다. 이제 준비되셨으면 시작해도 될까요?”

#42: Due to the ambiguity of the picture drawing of the distractor #4 ‘liver’, it was replaced with less ambiguous drawing of the same word.

#43: Due to unrecognizability of the picture drawing of the distractor #4, it was replaced with less ambiguous picture drawing of target word ‘honey’.

#45: The target item and corresponding stimulus were replaced with item #64 and its stimuli from the original K-BAT. This modification was made since four distractors of item #45 were not phonologically related to the target word and finding phonologically related distractors for the target word was impossible. Therefore, two syllable target word and its distractors with picture stimuli were borrowed from the original K-BAT to increase the difficulty of the test. Additional modification was made to the target word. Since the target word was not supposed to be included in the picture stimuli, original target word ‘고리’ which is represented in the stimulus #3 was replaced with ‘보리’ that is also phonologically related to the four distractors.

#46: The arrow of the picture drawing of the distractor #4 was modified in direction so that the arrow could point to the exact target item ‘foot’.

Syntactic Comprehension Subtest

#53: In order to increase the difficulty of the subtest, two additional items from the original K-BAT were added. Modification of stimulus book was not necessary since picture drawing of correct answers of the two added items were already included in the original stimulus page. Two added items are followings: ‘고양이가 개에게 물렸습니다’ and ‘고양이가 물고 있는 것은 개입니다’.

Series Subtest

#75 was replaced with #261 “숫자 1에서 25까지 순서대로 말해 주세요” from the original K-BAT in order to increase the difficulty of the test item. Additionally, since there are two ways to count in Korean, scoring criteria were expanded so that both ways of counting could be considered as correct answers.

The following subtests under the original short K-BAT are not modified and used entirely for the modified K-BAT: pointing, simplex and semi-complex commands, repetition of words and nonsense words, repetition of sentences, verbal fluency, semantic opposites, reading, copying, dictation, reading comprehension for words, and reading comprehension for sentences.

Appendix III. Scoring Procedures

K-BAT and E-BAT

Spontaneous Speech: Participants' spontaneous speech was recorded via audio recorder. The speech was scored on a scale from one to five for each of the six items: amount of speech, fluency, articulation, syntax, lexicon, and code-switching. The Maximum score of the subtest is 30.

Naming: Participants' verbal responses were recorded via audio recorder. Participants received 1 point for each correct answer for 10 items. Incorrect response received 0 point. A failure to give answer during 30 seconds time window was considered as no response and the participant received 0 point. The maximum score for the subtest is 10.

Pointing: Participants' physical hand gestures to point to or touch certain item on the table in front was used for scoring. Participants received 1 point for each correct pointing. Incorrect pointing received 0 point. A failure to point to certain item during 30 seconds time window was considered as no response and the participant received 0 point. The maximum score for the subtest is 5.

Simple and Semi-complex Commands: Participants' physical gestures were used for scoring. Participants received 1 point for each correct body or face gestures. Incorrect pointing received 0 point. A failure to make any response during 30 seconds time window was considered as no response and the participant received 0 point. The maximum score for the subtest is 6.

Complex Commands: Participants' physical gestures were used for scoring. Participants received 4 points for each item if all three commands were performed correctly in the right order. Participants received 3 points if all three commands were performed correctly, but in the wrong order. Participants received 2 points if two commands were performed correctly, irrespective of

the order. Participants received 1 point if only one command was performed correctly. Participant received 0 point if no command was performed correctly. A failure to make any response during 30 seconds time window was considered as no response and the participant received 0 point. The maximum score for the subtest is 12.

Verbal Auditory Discrimination: Participants' hand pointing to the provided picture stimuli was used for scoring. Participants received 1 point for each correct pointing to the target picture stimulus. Incorrect pointing received 0 point. A failure to make any response during 30 seconds time window was considered as no response and the participant received 0 point. The maximum score for the subtest is 7.

Syntactic Comprehension: Participants' hand pointing to the provided picture stimuli was used for scoring. Participants received 1 point for each correct pointing to the target picture stimulus. Incorrect pointing received 0 point. A failure to make any response during 30 seconds time window was considered as no response and the participant received 0 point. The maximum score for the subtest is 12.

Repetition of Words and Nonsense Words: Participants' verbal responses were recorded via audio recorder. Participants received 1 point for each item if the participant repeated exactly what he or she heard without any error. Participants received 0 if responses were incorrect. A failure to produce response during 5 seconds time window was considered as no response and participants received 0 point. The maximum score for the subtest is 12.

Repetition of Sentences: Participants' verbal responses was recorded via audio recorder. Participants received 1 point for each item if the participant repeated exactly what he or she heard without any error. Participants received 0 if responses were incorrect. A failure to produce

response during 5 seconds time window was considered as no response and participants received 0 point. The maximum score for the subtest is 3.

Series: Participants' verbal responses were recorded via audio recorder. Participants received 1 point if he/she performed the task perfectly without making any errors (i.e., omitting items, adding incorrect items, changing the order). Participants received 0 point if he/she made any errors. A failure to make any response during 30 seconds time window after promoting was provided was considered as no response and the participant received 0 point. The maximum score for the subtest is 2.

Verbal Fluency: Participants' verbal responses were recorded via audio recorder. Participants received 1 point for each correct names of animal. In addition, participants received 1 point if all the produced words were names of animals. Participants received 0 point if any one of the produced words were not names of animals. A failure to make any response during 30 seconds time window was considered as no response and the participant received 0 point.

Semantic Opposites: Participants' responses were recorded via audio recorder. The response was considered as correct and participants received 1 point if its meaning was opposite to, but not morphologically related to the stimulus words. Participants received 0 point if the response was incorrect or it was morphologically related to the stimulus word. The maximum score for the subtest is 5.

Reading: Participants' verbal responses were recorded via audio recorder. Participants received 1 point for each item if the participant read stimuli exactly without any error. Participants received 0 if responses were incorrect. A failure to produce response during 30 seconds time window was considered as no response and participants received 0 point. The maximum score for the subtest is 9.

Copying: Participants' written responses were used for scoring. Participants received 1 point for each correct response. Participants received 0 point if responses were incorrect. A failure to produce response during thirty seconds time window was considered as no response and participants received 0 point. The maximum score for the subtest is 2.

Dictation: Participants' written responses were used for scoring. Participants received 1 point for each correct response. Participants received 0 point if responses were incorrect. A failure to produce response during thirty seconds time window was considered as no response and participants received 0 point. The maximum score for the subtest is 3.

Reading Comprehension for Words: Participants' hand pointing to the provided picture stimuli were used for scoring. Participants received 1 point for each correct pointing to the target picture stimulus. Incorrect pointing received 0 point. A failure to make any response during 30 seconds time window was considered as no response and the participant received 0 point. The maximum score for the subtest is 4.

Reading Comprehension for Sentences: Scoring procedure for the subtest was same as the procedure used for the reading comprehension for words. The maximum score for the subtest is 4.

The total maximum score for each version of BAT, excluding the score from verbal fluency is 126. Scores for verbal fluency subtest were excluded in calculating maximum scores since the maximum score for the subtest could not be determined.

KETT

Word Recognition: Participants' verbal responses were recorded using audio recorder. Participants received 1 point for each correct response. Participants received 0 point if the

response was incorrect or if the participant failed to make any response within 5 seconds time window. The maximum score for the subtest is 5.

Translation of Words: Participants' verbal responses were recorded using audio recorder. Participants received 1 point for each correct response. If the participant produced different words from the provided correct answer, but acceptable, participant received 1 point. Participants received 0 point if the response was incorrect or if the participant failed to make any response within 5 seconds time window. The maximum score for the subtest is 10.

Translation of Sentences: Participants' verbal responses were recorded using audio recorder. Repetition of each stimulus was allowed for three times. The number of times participants requested repetition of the stimulus was noted, but did not affect the score. Participants received 3 points if all word groups were translated as suggested without any error, 2 points if 2 word groups were translated correctly without any error or 1 point if only one word group was translated correctly. A failure to produce response within 30 seconds time window was considered as no response and participants received 0 point. The maximum score for the subtest is 18.

Grammaticality Judgment: Participants' verbal responses were recorded using audio recorder. Participants received 1 point for each correct judgment (i.e., yes, it is grammatically correct sentence or no, it is not grammatically correct sentence). For the grammatically incorrect sentences, participants were asked to correct given sentences. Participants received additional 1 point for acceptably corrected sentences. Participants received 0 point if any error was present. If participants failed to make any response or to correct given grammatically incorrect sentences within 30 seconds time window, it was considered as no response and participants received 0 point. For four grammatically correct sentences (i.e., items #486, 492, 502 and 508), if

participants identified them correctly as grammatically acceptable sentences and did not attempt to correct them, participants received 2 point for each item. However, if participants identified them incorrectly as grammatically unacceptable sentences and corrected them with unacceptable sentences, participants received 0 point. If the participants corrected those sentences with acceptable sentences, participant received 1 point. The maximum score for the subtest is 16. The total maximum score for all subtests on each translation direction is 49.

Appendix IV. Korean-LexTALE

Korean word	Meaning
동단주	Nonword
자전거	Bicycle
다람쥐	Squirrel
슥버축	Nonword
옳다	Right
대폭	Drastically
권력	Authority
끔말	Nonword
넓다	Wide
어렵다	Difficult
아니다	Be not
몸	Body
학을험수	Nonword
맡다	Take care of
그런	Such
전화번호	Phone number
옆	Side
가렵다	Itchy
초닥	Nonword
끝	End
보다	See
솩	Nonword
알다	Know
두다	Put
모르다	Do not know, unaware
블서보	Nonword
마지막	Last
분위기	Atmosphere
갖다	Have
트합용초	Nonword
들어오다	Enter
돋다	Sprout
랭	Nonword
안타깝다	Regrettable
됄	Nonword
형닉석	Nonword
작다	Small
좋다	Good

서다	Stand
비졸	Nonword
기간	Period
중요하다	Important
경면촉	Nonword
일부	Part
환곱	Nonword
헙	Nonword
자신	Oneself
속	Inside
때문	Because
생각	Thought
알깁	Nonword
정궁벅	Nonword
시대	Era
버리다	Throw away
세계	World
외국	Foreign country
숟착	Nonword
끕	Nonword
중	Middle
종춘	Nonword
을학	Nonword
다리	Leg
잡다	Hold

References

Allen, M. J., & Yen, W. M. (2002). *Introduction to measurement theory*. Prospect Heights, IL: Waveland Press.

Anderson, E. C., Vanderhoff, A. M., & Donovick, P. J. (2013). A manifestation of the bilingual disadvantage in college-level writing. *International Journal Of Language Studies*, *7*(1), 139-150.

Benton, A. L., & Hamsher, K. D. (1994). Examen de Afasia Multilingue [Multilingual Aphasia Exam]. San Antonio, TX: The Psychological Corporation.

Bialystok, E., Craik, F.I.M., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. Journal of Experimental Psychology: Learning, Memory, and Cognition, 34, 859–873.

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

Bialystok, E., Craik, F. M., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science In The Public Interest*, *10*(3), 89-129.

Bitchener, J., Young, S., & Cameron, D. (2005). The Effect of Different Types of Corrective Feedback on ESL Student Writing. *Journal Of Second Language Writing*, *14*(3), 191-205.

Birdsong, D., Gertken, L. M., & Amengual, M. (2012). Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism. Retrieved 08 April, 2012, from COERLL, University of Texas at Austin https://sites.la.utexas.edu/bilingual/

Blumenfeld, H. K., Bobb, S. C., & Marian, V. (2016). The role of language proficiency, cognate status and word frequency in the assessment of Spanish–English bilinguals' verbal fluency. *International Journal Of Speech-Language Pathology*, *18*(2), 190-201.

Bosch, Karuse, & Leminen, A. (2017). The time-course of morphosyntactic and semantic priming in late bilinguals: A study of German adjectives. *Bilingualism: Language & Cognition*, *20*(3), 435-456.

Chung, H. H. (2006). Code Switching as a Communicative Strategy: A Case Study of Korean-English Bilinguals. *Bilingual Research Journal*, *30*(2), 293-307.

Gary, S. F., & Fennig, C. D. (2017). South Korea. Retrieved January 23, 2018, from https://www.ethnologue.com/country/KR

Fabbro, F. (1999). *The neurolinguistics of bilingualism : An introduction*. Hove: Psychology Press.

Fabbro, F. (2001). The bilingual brain: Bilingual aphasia. Brain and Language, 79, 201–210.

Fabbro, F., & Frau, G. (2001). Manifestations of aphasia in Friulian. Journal of Neurolinguistics, 14(2-4), 255-279.

Fabbro, F., & Paradis, M. (1995). Differential impairments in four multilingual clients with subcortical lesions. In M. Paradis (Ed.), Aspects of bilingual aphasia (pp. 139–176). New York: Elsevier.

Franzen, M. D. (2003). *Reliability and validity in neuropsychological assessment*. New York, NY: Kluwer Academic/Plenum.

Gertken, L. M., Amengual, M., & Birdsong, D. (2014). Assessing language dominance with the Bilingual Language Profile. In P. Leclercq, A. Edmonds & H. Hilton (Eds.), *Measuring L2 proficiency: Perspectives from SLA* (pp.208-225). Bristol: Multilingual Matters.

Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology, 16*(4), 562-576.

Gollan, T.H., Montoya, R.I., Fennema–Notestine, C., & Morris, S.K. (2005). Bilingualism affects picture naming but not picture classification. Memory & Cognition, 33, 1220–1234.

Gollan, T. H., Fennema-Notestine, C., Montoya, R. I., & Jernigan, T. L. (2007). The bilingual effect on Boston Naming Test performance. *Journal Of The International Neuropsychological Society*, *13*(2), 197-208.

Gómez Ruiz, I. (2008). Aplicabilidad del Test de la afasia para bilingües de Michel Paradis a la población catalano/ castellano parlante (Doctoral Thesis). Universidad de Barcelona.

Gomez-Ruiz, I. m., & Aguilar-Alonso, Á. (2011). Capacity of the Catalan and Spanish versions of the Bilingual Aphasia Test to distinguish between healthy aging, mild cognitive impairment and Alzheimer's disease. Clinical Linguistics & Phonetics, 25(6/7), 444-463.

Gray, T., & Kiran, S. (2013). A Theoretical Account of Lexical and Semantic Naming Deficits in Bilingual Aphasia. *Journal Of Speech, Language & Hearing Research*, *56*(4), 1314-1327.

Green, D. W. (2005). The neurocognition of recovery patterns in bilingual speakers with aphasias. In J. F. Kroll & M. B. de Groot (Eds.), Handbook of bilingualism: Psycholinguistic perspectives (pp. 516–530). New York: Oxford University Press.

Guion, S. G. (2005). Knowledge of English word stress patterns in early and late Korean-English bilinguals. *Studies In Second Language Acquisition*, *27*(4), 503-533.

von Hapsburg, D., & Bahng, J. (2006). Acceptance of background noise levels in bilingual (Korean-English) listeners. *Journal Of The American Academy Of Audiology*, *17*(9), 649-658.

Ha, S., Johnson, C. J., & Kuehn, D. P. (2009). Characteristics of Korean Phonology: Review, Tutorial, and Case Studies of Korean Children Speaking English. *Journal of Communication Disorders*, *42*(3), 163-179.

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production?. *Acta Psychologica*, *127*(2), 277-288.

Ivanova, M., & Hallowell, B. (2009). Short form of the bilingual aphasia test in Russian: psychometric data of persons with aphasia. *Aphasiology, 23*(5), 544-556.

Ivanova, M. V., & Hallowell, B. (2013). A tutorial on aphasia test development in any language: key substantive and psychometric considerations. *Aphasiology*, *27*(8), 891-920.

Juncos-Rabadán, O. (1994). The assessment of bilingualism in normal aging with the Bilingual Aphasia Test. Journal Of Neurolinguistics, 8(1), 67-73.

Jung, O. R. (2006). Daegu Aphasia Diagnostic Examination. Seoul: SigmaPress.

Kang, J. (2012). Do bilingual children possess better phonological awareness? Investigation of Korean monolingual and Korean-English bilingual children. *Reading & Writing*, *25*(2), 411-431.

Kang, J. Y. (2003). On the ability to tell good stories in another language: analysis of Korean EFL learners' oral "frog story" narratives. *Narrative Inquiry*, *13*(1), 127-149.

Kang, J. Y. (2005). Written Narratives as an Index of L2 Competence in Korean EFL Learners. *Journal Of Second Language Writing*, *14*(4), 259-279.

Kang, K., & Guion, S. G. (2006). Phonological Systems in Bilinguals: Age of Learning Effects on the Stop Consonant Systems of Korean-English Bilinguals. *Journal Of The Acoustical Society Of America*, *119*(3), 1672-1683.

Kendall, D., Edmonds, L., Van Zyl, A., Odendaal, I., Stein, M., & van der Merwe, A. (2015). What can speech production errors tell us about cross-linguistic processing in bilingual aphasia? Evidence from four English/Afrikaans bilingual individuals with aphasia. *The South African Journal Of Communication Disorders*, *62*(1), E1-E10.

Kim, H., & Na, D. L. (1999). Normative data on the Korean version of the Boston Naming Test. Journal of Clinical And Experimental Neuropsychology, 21(1), 127-133.

Kim, H., & Na, D. (2001). Paradise Korean version- Western Aphasia Battery (P K-WAB). Seoul: Paradise Institute for Children with Disabilities.

Kim, H & Na, D. L. (2004). Normative Data on the Korean Version of the Western Aphasia Battery. Journal Of Clinical & Experimental Neuropsychology, 26(8), 1011-1020.

Kim, J. (2010). Perception and production of English front vowels by Korean speakers. *Phonetics and Speech Sciences, 2*(1), 51-58.

Kim, Y. (2009). Crosslinguistic Influence on Phonological Awareness for Korean-English Bilingual Children. *Reading And Writing: An Interdisciplinary Journal*, *22*(7), 843-861.

Ko, H., Ionin, T., & Wexler, K. (2009). L2 acquisition of English articles by Korean speakers. *Handbook of East Asian Psycholinguistics: Korean, Cambridge University Press*, 286-304.

Kohnert, K.J., Hernandez, A.E., & Bates, E. (1998). Bilingual performance on the Boston Naming Test: Preliminary norms in Spanish and English. Brain and Language, 65, 422–440.

Koumanidi Knoph, M. I. (2011). Language assessment of a Farsi--Norwegian bilingual speaker with aphasia. Clinical Linguistics & Phonetics, 25(6/7), 530-539.

Krishnan, G., & Mathew, R. E. (2017). Short version of the Bilingual Aphasia Test in Malayalam. Annals of Indian Academy Of Neurology, 20(3), 217-219.

Lee, H. J., Kim, Y. T., & Yim, D. (2013). Non-word repetition performance in Korean-English bilingual children. *International Journal Of Speech-Language Pathology*, *15*(4), 375-382.

Lee, K. J., Lee, C. J., Kim, J. H., Jung, S. M., & Hwang, S. J. (2009). The Concurrent Validity of the 'Korean Test for the Differential Diagnosis of Aphasia'. Korean Journal of Communication Disorders, 58-69. Retrieved November 3, 2017.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343.

Lorenzen, B., & Murray, L. L. (2008). Bilingual Aphasia: A Theoretical and Clinical Review. American Journal of Speech-Language Pathology, 17(3), 299-317.

Manuel-Dupont, S., Ardila, A., Rosselli, M., & Puente, A. E. (1992). Bilingualism. In R. McCaffrey & A. E. Puente (Eds.), Handbook of neuropsychological assessment: A biopsychosocial perspective (pp. 193–210). New York: Plenum Press.

Marrero, M. Z., Golden, C. J., & Espe-Pfeifer, P. (2002). Bilingualism, brain injury, and recovery: Implications for understanding the bilingual and for therapy. *Clinical Psychology Review*, *22*(3), 463-478.

Mayo, L. H., Florentine, M., & Buus, S. (1997). Age of second language acquisition and perception of speech in noise. *Journal of Speech, Language and Hearing Research, 40,* 686–693.

Mindt, M. R., Arentoft, A., Germano, K. K., D'Aquila, E., Scheiner, D., Pizzirusso, M., & ... Gollan, T. H. (2008). Neuropsychological, cognitive, and theoretical considerations for evaluation of bilingual individuals. Neuropsychology Review, 18(3), 255-268.

Muñoz, M. L., & Marquardt, T. P. (2003). Picture naming and identification in bilingual speakers of Spanish and English with and without aphasia. *Aphasiology*, *17*(12), 1115-1132.

Muñoz, M., & Marquardt, T. (2008). The performance of neurologically normal bilingual speakers of Spanish and English on the short version of the bilingual aphasia test. *Aphasiology, 22*(1), 3-19.

Paradis, M. (1977). Bilingualism and aphasia. In H. Whitaker & H. Whitaker (Eds.), Studies in neurolinguistics (Vol. 3, pp. 65–121). New York: Academic Press.

Paradis, M., Goldblum, M., & Abidi, R. (1982). Alternate antagonism with paradoxical translation behavior in two bilingual speakers with aphasia clients. Brain and Language, 15, 55–69.

Paradis, M., & Libben, G. (1987). *The assessment of bilingual aphasia*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Paradis, M. (2001). Bilingual and polyglot aphasia. In R. S. Berndt (Ed.), Handbook of neuropsychology (2nd ed., pp. 69–94). Oxford: Elsevier Science Ltd.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism* (Studies in bilingualism, v. 18). Amsterdam: J. Benjamins Pub.

Paradis, M. (2011). Principles underlying the bilingual aphasia test (BAT) and its uses. *Clinical Linguistics & Phonetics, 25*(6-7), 427-443.

Park, H. S. (2005). Development and standardization of Korean Test for Differential Diagnosis of Aphasia and characteristics of the Korean aphasics. Proceedings of the Korean Academy of Speech-Language Pathology and Audiology annual conference. Cheonbuk University, Jeonju.

Park, T. (2005). English article use by advanced Korean EFL learners. *The Korea Association of Foreign Languages Education, 12*(1), 35-59.

Pepe, M. S. (2003). The statistical evaluation of medical tests for classification and prediction. New York: Oxford University Press.

Peristeri, E., & Tsapkini, K. (2011). A comparison of the BAT and BDAE-SF batteries in

determining the linguistic ability in Greek-speaking patients with Broca's aphasia. Clinical Linguistics and Phonetics, 25.

Pindzola, R. H., Plexico, L. W., & Haynes, W. O. (2016). *Diagnosis and evaluation in speech pathology* (9th ed.). Boston: Pearson.

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

Rey, G. J., Sivan, A. B., & Benton, A. L. (1991). Multilingual Aphasia Examination–Spanish Version. Lutz, FL: Psychological Assessment Resources.

Roberts, P.M., Garcia, L.J., Desrochers, A., & Hernandez, D. (2002). English performance of proficient bilingual adults on the Boston Naming Test. Aphasiology, 16, 635–645.

Roberts, P. M. (2008). Aphasia assessment and treatment for bilingual and culturally diverse patients. In R. Chapey (Ed.), Language intervention strategies in aphasia and related neurogenic communication disorders (5th ed., pp. 245– 275). Baltimore, MD: Lippincott Williams & Wilkings.

Ross, K. B., &Wertz, R. T. (2004). Accuracy of formal tests for diagnosing mild aphasia: An application of evidence-based medicine. Aphasiology, 18, 337–355.

Rosselli, M., Ardila, A., Araujo, K. Weekes, V.A., Caracciolo, V. Padilla, M., & Ostrosky–Solis, F. (2000). Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. Applied Neuropsychology, 7, 17–24.

Rosselli, M., Ardila, A., Salvatierra, J., Marquez, M., Matos, L., & Weeks, V. A. (2002). A cross-linguistic comparison of verbal fluency tests. *International Journal Of Neuroscience*, *112*(6), 759.

Sadat, J., Martin, C. D., Alario, F. X., & Costa, A. (2012). Characterizing the bilingual disadvantage in noun phrase production. *Journal Of Psycholinguistic Research*, *41*(3), 159-179.

Shi, L. F. (2010). Perception of acoustically degraded sentences in bilingual listeners who differ in age of English acquisition. *Journal of Speech, Language and Hearing Research, 53,* 821–835.

Strauss, E., Sherman, E. M. S., & Spreen, O. (with Slick, D. J.) (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). Oxford: University Press.

Suh, E. (2017). Incomplete acquisition in the nominal domain of Korean by heritage language speakers. *Dissertation Abstracts International*, *77*(7),

Swaminathan, H., & Rogers, H. J. (1990). Detecting Differential Item Functioning Using Logistic Regression Procedures. *Journal of Educational Measurement, 27*(4), 361-370.

Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Flege, J. (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal Of Phonetics*, *33*(3), 263-290.

US Census Bureau (2014). *American Community Survey*. Retrieved December 2017.

US Census Bureau (2011). *Language Use in the United States*. Retrieved December 2017.

Wang, M., Park, Y., & Lee, K. R. (2006). Korean-English biliteracy acquisition: cross-language phonological and orthographic transfer. *Journal Of Educational Psychology*, *98*(1), 148-158.

Yeni-Komshian, G. H., Flege, J. E., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean-English bilinguals. *Bilingualism: Language And Cognition*, *3*(2), 131-149.

Yoo, J. y., & Kaushanskaya, M. (2012). Phonological memory in bilinguals and monolinguals. *Memory & Cognition*, *40*(8), 1314-1330.