

# **How Native Language Affects Functionally Relevant Item Selection**

by

Lissien Esther Barahona Sanchez

A thesis submitted to the Graduate Faculty of

Auburn University

in partial fulfillment of the

requirements for the Degree of

Master of Science in Communication Disorders

Auburn, Alabama

May 1<sup>st</sup>, 2021

Key Words: functionally relevant items, Spanish, English, bilingual, blank Canvas, language

Copyright 2021 by Lissien Esther Barahona Sanchez

Approved by

Dr. Dallin J. Bailey, Chair, Assistant Professor, Department of Speech and Hearing Sciences

Dr. Nancy Haak, Associate Professor, Department of Speech and Hearing Sciences

Dr. Megan- Brette Hamilton, Assistant Professor, Department of Speech and Hearing Sciences

Dr. Marisha Speights Atkins, Assistant Professor, Department of Speech and Hearing Sciences

## Abstract

**PURPOSE:** A challenge speech language pathologists face when establishing a therapy plan to rehabilitate a significant loss of language and communication ability is the identification and selection of functionally relevant items to target during therapy. There are many variables which must be considered when selecting functionally relevant items, including the native language of the patient. Accounting for the differences which arise from native language could improve the quality of the selected target lists and improve the quality of therapy. The overall purpose of this study is to determine how native language impacts the selection of functionally relevant items. A secondary aim is to determine how corpus type affects the selection of functionally relevant items.

**METHODS:** Twenty-three Native Spanish speakers, eighty-one native English speakers and fifty-nine bilingual speakers responded to blank canvas and open-ended questions. Their responses were gathered and used for the creation of six different corpora. The effect of language group and corpus type was analyzed descriptively and inferentially.

**RESULTS:** Using descriptive statistics, multivariate analysis, test of between subjects and chi square analysis it was determined that both language groups and corpus type had significant effects on the psycholinguistic variables of concreteness ratings, word length in phoneme, part of speech occurrence, and external corpus frequency.

**CONCLUSIONS:** The results display significant effects of native language on the selection of functionally relevant items. Understanding these differences may increase the accuracy and effectiveness of target lists during therapy for individuals with different languages.

## Acknowledgments

I would first like to thank my mentor Dr. Bailey for the constant encouragement, guidance, and support throughout this process. I have had the privilege of having him as my professor, mentor, and supervisor throughout my graduate career. Dr. Bailey offered his help to me when I was completing my undergraduate thesis, even before I was his student. That is simply the kind of professor he is, the kind that will go the extra mile to help a student, any student. Thank you, Dr. Bailey for teaching me everything I know about motor speech disorders, for helping me become a better researcher and for teaching to love my profession and enjoy every step of the arduous process.

Thank you to my committee members, Dr. Haak, Dr. Hamilton and Dr. Speights Atkins, for agreeing to take this journey with Dr. Bailey and me. Thank you for providing resources, feedback, and statistical help. You all played an integral part in the success of this project.

I'd like to also thank my friends who have been a constant encouragement throughout my academic journey. They have been by my side through long study nights and encouraged me on days in which my academic burdens seem far too great. Through encouraging texts, coffee runs and study breaks, you allowed me to see that I was not alone in completing this journey.

Thank you also to my family. Papá, por siempre recordarme que soy capaz de lograr todo lo que me ponga en mente, y recordarme que mi mayor esfuerzo es el mejor esfuerzo. Y por todas las noches de estudio y explicaciones de estadística. Mamá, por siempre estar a mi lado cuando estaba a punto de rendirme. Por ser mi mayor fan y mi amiga mas honesta, se que siempre puedo contar con tu apoyo. A mis hermanos, por sus consejos, apoyo técnico, y

constantes palabras de aliento. Sin mi familia no podría haber logrado llegar a la culminación de mi maestría.

Finally, I'd like to thank my fiancé, Hayden. Thank you for always being there for me. Thank you for patiently listening to my hopes, goals and doubts. Thank you for always being willing to help me in any way to further my research, whether it be with excel sheets, loaned laptops, car study sessions, or table reformatting. You were a tremendous help throughout this process. I truly could not have done this without your support. You have always believed in me and I am so thankful for that.

## Table of Contents

Abstract.....	ii
Acknowledgments.....	iii
List of Tables .....	vii
List of Abbreviations .....	viii
Definitions.....	ix
Chapter 1: Review of Literature .....	1
The Need for Functionally Relevant Items.....	1
Current Practices of Vocabulary Selection and its Challenges.....	3
Important Variables in Functionally Relevant Vocabulary Selection .....	5
Purpose.....	12
Chapter 2: Methods.....	13
Participants: .....	13
Recruitment Strategy .....	13
Survey Development and Administration.....	12
Corpus Construction .....	15
Data Analysis .....	15

Chapter 3: Results .....	17
Participants.....	18
Descriptive Statistics .....	20
Inferential Statistics .....	22
Chapter 4: Discussion .....	25
Conclusion .....	33
References .....	32
Appendix A: Survey Questions .....	40
Appendix B : Lists of Top 50 Most Frequent Words in Each Corpus.....	49
Appendix C : Lists of Top 10 Most Frequent Nouns and Verbs in Each Corpus.....	61

## List of Tables

Table 1: Participant Demographics.....	17
Table 2: Open Ended Question Results .....	19
Table 3: Blank Canvas Results .....	20

## List of Abbreviations

MEBC	Monolingual English Blank Canvas
MEOEQ	Monolingual English Open-Ended Question
MSBC	Monolingual Spanish Blank Canvas
MSOEQ	Monolingual Spanish Open-Ended Question
BBC	Bilingual Blank Canvas
BOEQ	Bilingual Open-Ended Question
WLIP	Word Length in Phonemes
POS	Part of Speech



## Definitions

Corpus: A collection of words for analysis

Tokens: Total number of words in a work or a corpus

## Chapter 1

### Introduction

#### **The Need for Functionally Relevant Items**

In aphasia, word-finding deficits are one of the most common characteristics (Goodglass & Wingfield, 1997; Raymer, 2005). Given that word finding deficits are such a common defining characteristic of aphasia, aphasiologists are motivated to find relevant, functional, and useful vocabulary to work on during therapy (Renvall, Nickels, & Davidson, 2013). Word finding meta-analyses have revealed that the strongest advances achieved in anomia therapy are those which are practiced during therapy sessions (Wisenburn & Mahoney, 2009; Kim et al., 2020). Regardless of the manner of exposure (whether it be semantically, phonologically or through mixed therapies), research shows that exposed words display the most gains (Wisenburn & Mahoney, 2009). Given that words practiced during therapy sessions show the most gain, it is important to ensure that those words are functionally relevant.

Renvall et al. states that functional vocabulary can be divided “into two categories — personally chosen vocabulary and generally frequent vocabulary” (2013, p. 636). Renvall et al. describes personally chosen items as “items a person with aphasia, his/her significant other(s) and/or their clinician, identify as important for their communicative success and should be targeted in a treatment program tailored for that particular individual” (Renvall et al., 2013b, pg. 636). Items that may be found under the category of personally chosen items could be considered topics of interest to that person, or items which a familiar communicator believes that person would want to continue to use. Generally frequent items are defined as “those that unimpaired adult speakers use frequently in their everyday communication” (Renvall et al., 2013b pg. 636). Items may be found under the category of generally frequent items are those which have been

identified through “objective counts of word frequency from large samples of spoken language” (Renvall et al., 2013b, pg. 636). Due to the nature of their selection, generally frequent items are considered functionally relevant for a population rather than an individual (Renvall et al., 2013b).

A similar way of defining functional vocabulary is observed in augmentative and alternative communication (AAC). In the selection of functionally relevant vocabulary, AAC facilitators divide their lexicon selection into two categories: core and fringe vocabulary. In 1988, Yorkston et al. defined core vocabularies as small vocabularies that have minimal change over time and show a high degree of similarity across differentially normed users (Yorkston et al., 1988). On the other hand, fringe vocabularies are more individualistic and consist mainly of content words (Balandin & Iacono, 1998). These two definitions can be compared to personally relevant items and generally frequent items. Core vocabulary would be similar to generally frequent items in the sense that these items are normed throughout a variety of users and personally relevant items resemble fringe vocabulary in their individualistic nature.

A strong link between words practiced during therapy sessions and greater generalization of said words, is in fact a concept that is valid throughout language and vocabulary recovery as a whole and is not exclusive to anomia treatment during aphasia therapy. A comparable adherence to the concept of greater generalization linked to practice during therapy can be seen in the selection of vocabulary for the development, maintenance, and rehabilitation of language in individuals who communicate through the use of AAC (Beukelman et al., 1991). In fact, the early stages of vocabulary selection research focused on the core and fringe vocabulary used by users of augmented communication (Farrier et al., 1985).

## **Current Practices of Vocabulary Selection and its Challenges:**

Current practices in speech language pathology for the selection of functionally relevant vocabulary are comprised of a selection of generally frequent items and the identification of personally chosen items (Renvall et al., 2013b). Generally frequent items are commonly identified using a language corpus based on the idea of objective counts and word usage in spoken language (Renvall et al., 2013b). This is done by searching for the most frequently used words in a language using a language corpus, such as the British National Corpus (Renvall et al., 2013a ). Personally chosen items are generally identified by carrying out assessments and interviews with family members in which they express their communicative priorities (Renvall et al., 2013a).

The selection of functionally relevant items brings forth many challenges. The first challenge surfaces through the use of language corpora. Although useful, a language corpus is mainly based on the occurrences of words in written language, which may not account for language used in everyday conversations (Brysbaert et al., 2014). The second challenge arises when selecting personally chosen items. There are not many materials designed to target the specific selection of vocabulary (Renvall et al., 2013b). Therefore, speech language pathologists rely on the use of interviews in which they ask the individual with the language deficit and their communication partners what they find personally relevant. A common approach is the “blank page” technique as it is named by Renvall et al. (2013). Renvall et al. state that one issue with the blank canvas technique is that it often overwhelms the patient and their communication partners, yielding an inaccurate and noun heavy list of desired words (2013). This generally leads to a very concrete, noun heavy lexicons (Renvall et al., 2013b). A noun heavy lexicon is not representative of the everyday spoken language. Conversations require verbs, adjectives, articles,

words of high concreteness and words of low concreteness. Therefore, a noun heavy, high concreteness list of words is simply not representative of the diverse lexicon that one can observe in everyday conversations and quantify using a language corpus.

### **Problems with the Blank Canvas Approach**

Often, treatment lists target high concreteness nouns (Bailey, 2020). However, lists which are mainly comprised of high concreteness nouns limit the communication needs of an individual who has a severe language impairment (Palmer et al., 2017). Palmer et al asked 100 people with aphasia to identify 100 words that would be primarily important for them to be able to say. With help of their communicative partners and 18 topic pictures the participants developed their lists. The resulting lists and analysis revealed that most of the words selected could be related to major factors such as the ones described by Worrall's research published in 2011. These factors include the desire to return to pre-stroke life, partake in important relationships, and be respected by highlighting their pre- morbid skills and accomplishments. However, the words from Palmer et al. did not necessarily adequately address the lexical requirements for Worrall's categories.

In 2011 Worrall noted that services are increasingly influenced by the International Classification of Functioning, Disability and Health (ICF), however it is not yet known whether this framework truly encompasses the goals of people with aphasia (Worrall et al., 2011). Worrall et al. conducted semi-structured in-depth interviews with 50 participants with aphasia post-stroke, which were videotaped and transcribed verbatim (2011). The interviews were then analyzed using qualitative content analysis in order to describe goals of people with aphasia according to the ICF. The results yielded nine categories that described the needs of aphasia patients as follows:

Participants with aphasia wanted to return to their pre-stroke life and to communicate not only their basic needs but also their opinions. They also wanted information about aphasia, stroke, and available services; more speech therapy; greater autonomy; and dignity and respect. They identified the importance of engagement in social, leisure, and work activities as well as regaining their physical health. Interestingly, their goals included wanting to help others (Worrall et al., 2011)

The categories demonstrated that aphasia patients' goals could be linked to all ICF components within the ICF (Worrall et al., 2011). This connection between the goals of aphasia patients and the ICF framework allows us to see that the selection of communicative items must account for the desires of patients to be more than a sustained member of society but truly incorporated and contributing.

Palmer et al.'s study showed that patients selected words in topics such as food and drink, nature and gardening, entertainment, place, people, house, clothes, and travel (Palmer et al., 2017). Different parts of speech, frequencies in use, and concreteness were observed. Further efforts must be made to include these variations and even further considerations must be made to understand which factors could affect these variable factors themselves.

### **Important Variables in Functionally Relevant Vocabulary Selection**

Overall, the challenge that is observed in obtaining both generally frequent items and personally chosen items is the difficulty of accounting for the various sources of variability in frequently used words. Speech therapists struggle to maintain an accurate and realistic balance of varied parts of speech and levels of concreteness (Bailey, 2016). A second observed challenge is accounting for the variability that will undeniably arise from individual preferences while maintaining a functional, flexible and specific list of words for therapy.

A realistic and functional lexicon should be diverse in its part of speech variety. Research by Boenisch & Soto shows that core vocabulary,

includes a variety of different word classes: pronouns (e.g., I, me, you, him), verbs (e.g., go, want, put, get, let), auxiliary verbs (e.g., is, do, have, can, could, will), adjectives (e.g., good, bad), adverbs (e.g., again, now, here, there, more), prepositions (e.g., in, on, with, of, for), determiners (e.g., this, that), conjunctions (e.g., and, or, because), interjections (e.g., yes, no, please, sorry), question words (e.g., who, what, when, where, why), and nouns (e.g., house, tree, boy) (2015, pg. 77).

Lexicons selected for therapy should reflect this diversity in order to foment greater generalization of practiced targets. Given the nature of generalization it is important to select words with linguistic properties that will aid sentence formation and yield more generalization. This is based on the idea that “underlying, abstract, properties of language will allow for effective generalization to untrained structures that share similar linguistic properties” (Thompson & Shapiro, 2005, pg. 1021). Grammatical roles, semantic fields, syntactic properties are among the many categories in which word-classes differ (Milman et al., 2014). These differences can even be observed at the neurological level. For example, when comparing verbs and nouns we can see that they differ in word meaning, as verbs are action driven. Neurologically, we see that brain damage closer to regions of motor planning result in a greater impairment of verbs, because the action- object distinction between verbs and nouns cause a difference at semantic and neurological levels (Vigliocco, Vinson, Druks, Barber, & Cappa, 2011). Nouns and verbs are two very important parts of speech to consider during production. When selecting vocabulary for therapy, speech language pathologists tend to include these parts of speech as they are useful carriers of main ideas within sentences. However, one does not often

encounter many adjectives within a word list to target during therapy, even though research suggests that adjectives possess a combination of associative, semantic, syntactic, morphological, and affective relations, which provide a priming effect that facilitate language production. (Milman et al., 2014). The effect of the part of speech and its psycholinguistic properties must be considered when selecting vocabulary, as it is evident that said properties will be key players in the generalization of skills acquired during therapy.

Yet another important factor in the area of generalization is concreteness. Concrete words have high imageability and can be experienced through senses (Sandberg & Gray, 2020). The concreteness of words affects the performance of said word in unimpaired adults in what is known as the concreteness effect (Sandberg & Gray, 2020). This tells us that there are different representations and processing for high concreteness words as opposed to low concreteness, “abstract” words (Sandberg & Gray, 2020). It should be stated that “training abstract words in a particular context-category promotes generalization to concrete words but not vice versa” (Sandberg & Kiran, 2014, p. 738). Therefore, ensuring a fair inclusion of abstract words in a list of functionally relevant words for treatment will likely ensure higher generalization of vocabulary all together. Given that the use of abstract words varies due linguistic experience it is important to consider native language when selecting vocabulary for therapy.

### **The Role of Native Language**

Previously it was stated that two major challenges that arise when selecting functionally relevant items for therapy are capturing the true diversity of different parts of speech and concreteness levels in frequently used language and capturing individual variability. It is not yet known how native language affects functionally relevant vocabulary selection. It is likely that native language may impact functionally relevant vocabulary selection in many ways, and these



differences can be separated into general differences between particular languages (language differences) and differences between individual speakers (speaker differences).

### **Language Differences**

In this study we will focus on the linguistic differences of Spanish and English. Spanish and English have many linguistic variabilities in every area of linguistics, including syntax, phonology and semantics. Perhaps the most obviously different domain amongst languages is phonology. Languages assign different phonemes to their mental representations in order to form words. This can pose a challenge for language learners because at the phonetic level, two languages may exhibit subtle variations which adult second language (L2) learners are often not as sensitive to (Imai et al., 2005). Spanish and English differ in the specific phonemes used, as well as the phonology of the words.

Languages also differ according to syntax. For example, when talking about periphrastic constructions vs inflections it has been noted that bilingual speakers who borrow and code switch between Spanish and English give etymologically English verbs Spanish tense/aspect and subject-agreement inflection but English adjectives are never inflected for gender and number (Pfaff, 1979). An additional difference is the marking of (grammatical) gender in nouns which is mandatory in Spanish but occurs less in English (Arbesman et al., 2010).

Finally, languages also differ according to semantics. Charteris-Black and Ennis (2001) studied use of metaphor in Spanish and English financial reporting. They looked at 203 financial reports of published in both languages regarding the 1997 stock market crash. The study revealed that there were similarities “in conceptual and linguistic metaphors between the two languages, but some differences in the frequency of particular linguistic metaphors” (Charteris-Black & Ennis, 2001, pg. 252). The study revealed that both languages share conceptualizations of the

economy as an organism, and they both view the market movements as physical movements and decreases in market as natural disasters (Charteris-Black & Ennis, 2001). However, the study showed that “while in Spanish reporting there is a preference for metaphors based on psychological mood and personality [such as “sufrir” (suffering), “temor” (fear)], in English reporting there is a higher frequency of nautically based metaphors [“such as plunge”, “weather the storm”]” (Charteris-Black & Ennis, 2001, pg. 262). The differences observed in the metaphors used in the different languages shows that cultural experiences and native language influence the selection of lexical items to confer an idea.

It is important to understand linguistic differences in languages since their impact goes beyond word production, but deeper, into the semantic mapping, mental representations and word retrievals that speech language pathologist attempt to repair post impairment through the use of word lists of functionally relevant items.

### **Speaker Differences**

Topic selection can also create variability when selecting functionally relevant items for a word list. However, unlike parts of speech and concreteness, this difference is more speaker dependent. This shift in difference causation is a result of differences among speakers. Speakers may differ according to topic. Topic can be affected by many aspects including when the speaker chooses to have a conversation, level of familiarity between communicators and satisfaction with the topic (Arimoto et al., 2019). Two additional causes for differences among speakers are age and native language. However, these differences do not create isolated effects; in fact they can further influence the general differences of parts of speech and concreteness as well as the speaker difference of topic selection.

### ***Age***

As a person ages their roles change. Adults desire to have access to higher education which prepares them for economic and civic roles of adulthood (Bryen, 2008). Adults desire to have productive roles within their society, such as the roles of romantic partner, workers and educated individuals (Bryen, 2008). The sufficiency of vocabulary to support a variety of adult roles is a current concern within professionals who habilitate and rehabilitate communication. The lexicon selected for adults should enable them to accomplish these roles.

Older adults also have a generational responsibility for transmitting cultural values and continuity with the past (Albert & Trommsdorff, 2014). Therefore, vocabulary selection for adults must include words that allow them to communicate time and opinions (Stuart et al., 1997). Their linguistic experience is more sophisticated; therefore, their lexicon should reflect that.

### *Native Language*

Native language is an integral component of cultural experience and is impacted by the culture of and region in which it is acquired. Spanish L1 speakers who developed their language within the United States hold within their spoken language examples of how region impacts native language. According to research, Spanish speakers utilize English lexical items both assimilated or unmodified (Lipsky, 1986). As opposed to a Spanish speaker born within a primarily Spanish speaking country, a Spanish speaker in a majority English-speaking country may use the phrase “te lo doy pa tras” when indicating that she “I will give it back to you” (Lipsky, 1986). On the other hand a Spanish speaker born in a Spanish-speaking region might use the phrase “te lo regreso” (Lipsky, 1986). The first example shows how the Spanish speaker who acquired their language within the United States borrowed English lexical items “back” with the determiner “to” with the aid of you and translated it to Spanish instead of using a gendered

determining suffix like a Spanish speaker from a Spanish speaking region likely would. Regional differences are observed within the same language among different regions. A 2007 study suggests that significant differences were found during semantic verbal fluency tasks among Spanish speakers from different countries (Ostrosky-Solis et al., 2007). However, the tasks in the study were not presented in the same way to all participants. Therefore, one cannot conclude that the differences observed due to cultural differences warrant a different assessment or task for each region. Although the isolated cultural difference is not enough to warrant much differentiation in tasks or assessment, the difference itself does suggest that more research must be done to understand the effects of region, culture and native language on the linguistic system.

## **Bilingualism**

In order to understand the role of a functionally relevant items in the process of language rehabilitation we must first understand the general organization of a bilingual brain and the generalization patterns of bilingual aphasia therapy. According to research a bilingual language system is not two unilingual language systems added together, but an integration of two languages into a single language system (Ansaldi et al., 2008). At a neuroanatomical level it is most probable that two languages are represented as different microanatomical systems within the same cerebral regions, creating two linguistic subsystems (Gomez- Ruiz, 2009). Although the languages are represented within one region (such as the left perisylvian region), in different subsystems, neuroanatomical differences alone, do not account for variability of recovery between languages in patients who have lost their ability to communicate. The languages' performance may vary due to frequency of use, proficiency and setting of acquisition (Gomez-Ruiz, 2009). When dealing with bilingual aphasia recovery it is important to consider that either language could recover quicker than the other or they could recover in a parallel pattern. A closer

look at pattern of acquisitions shows that the context of acquisition whether procedural or declarative would play a role in recovery patterns. L1s are usually acquired through procedural memory, whereas L2s (especially at an older age) are usually acquired through declarative memory which is associated to metalinguistic knowledge of lexicon and grammar (Gomez-Ruiz, 2009). Therefore, when new pathways are attempted to be created through therapy, it is possible that the language acquired through the declarative memory would begin to improve more quickly. Great consideration should be given to the differences, in recovery and how the usually lexical and grammatical approach of therapy would impact the recovery of one language over the other one. It is important to pay close attention to the grammatical weight of the functionally relevant items selected in order to achieve recovery of language, as well recovery of the most functional language.

### **Purpose**

- The first aim of this study is to determine whether native language (Spanish or English) and bilingual status impacts the selection of functionally relevant words.
- A second aim of this study is to create corpora from the responses provided from participants in order to observe and analyze the different frequency of words and parts of speech within their responses to determine the variabilities which native language influences in lexicon selection.
- A third aim of the study is to compare the words from this study's corpus to a larger world corpus of both the Spanish and English language.
- The final purpose of this study is to compare the parts of speech occurrence of open-ended questions vs. blank canvas questions.

## Chapter 2

### **Methods**

#### **Participants**

Twenty-three native Spanish speakers, eighty-one native English speakers and fifty nine bilingual speakers (both concurrent speakers and simultaneous speakers) whose language and cognitive abilities were within normal limits, were recruited for the study. Native language was defined as the first language spoken by an individual. Bilingual speakers were required to use their L2 at a minimum of 2 times per month. Bilingual speakers were asked to report their pattern of acquisition and self-rate their level of proficiency in both languages. All participants reported to have normal cognitive functions and abilities. All participants were also asked to report their country of origin and residence. All participants were over 19 years of age and reported completing at least 10 years of education. All participants reported normal or corrected-to-normal hearing and visual acuity. All participants reported having no previous language delays or impairments or have any current concerns regarding their language. Participants gave informed consent to participate in the study, which was approved by the Auburn University Institutional Review Board.

#### **Recruitment Strategy**

To recruit participants, notices regarding the study were shared and posted online in different social media platforms, including over 30 ASHA and Facebook special interest groups. Prospective participants then accessed the survey through the shared link and completed it.

#### **Survey Development and Administration**

The first question in the survey was designed to be a “blank canvas” question in which the participant was asked to provide twenty five words which they deemed necessary and

functionally relevant. This “blank canvas” format of interview was included in all the surveys in order to reflect current practices of vocabulary selection. The remainder of the survey was designed to elicit open-ended responses to specific functional situations. Topical and specific questions were developed using components of the International Classification of Functioning Disability and Health framework which was set in place to ensure the mental and social functioning of a person who is afflicted by a disease or disability. Activity, participation, and environmental factors were used as basis to develop questions for the survey. Furthermore, using the topic table developed by Renvall et al., (2013). subtopics were selected to develop questions within the categories presented by the ICF framework. To avoid noun bias in participants responses the questions were designed to elicit dialogue rich responses in both concrete and abstract verbs, nouns, pronouns, and adjectives which participants might use during their daily activities, conversations and within their environment. The questions were tested with a small number of subjects (5 people) to determine their ease of comprehension (as qualitatively rated by each participant) and expected response. Following the trial, the questions were uploaded to Qualtrics, an online surveying platform through which the survey was shared to obtain results for this study. The survey questions may be found in Appendix A. Participants responded to questions using the online surveying platform. Native language speakers were presented the questions in sets of five with the option of opting out of the survey after responding to 5 questions. Bilingual speakers were presented the questions in sets of 6 in which 3 questions were asked both in Spanish and English separately with the option of opting out of the survey after responding to 6 questions. The results were gathered and analyzed.

## **Corpus Construction**

An online corpus builder (Sketch Engine) was used to analyze the text responses and identify high frequency words. Each answer provided by the participants was converted to a singular text file and uploaded into the corpus. Six corpora were developed--a Monolingual English Blank Canvas corpus (MEBC), a Monolingual English Open Ended Question corpus (MEOEQ), a Monolingual Spanish Blank Canvas corpus (MSBC), a Monolingual Spanish Open Ended Question corpus (MSOEQ), a Bilingual Blank Canvas corpus (BBC) and a Bilingual Open Ended Question corpus (BOEQ). The 50 most frequent tokens within the sample, excluding punctuation, were then extracted from each corpus. Each corpus was set up to display the words by frequency, word, lemma, and part of speech. The corpus builder also showed the internal corpus frequency of each word within each corpus. The frequency and concreteness results from this corpus were then compared to results presented by two corpora for each individual language: for the English, the SUBTLEX-US corpus was used (Brysbaert New, 2009); for the Spanish, the SUBTLEX-ESP was used (Cuetos et al., 2011) These resources provide data that allows looking up specific words and receiving the external concreteness and frequency ratings

## **Data Analysis**

To begin the analysis, the participants' responses were gathered. Following the selection of data that met eligibility criteria, the responses were entered into word documents where they were standardized for analysis. The standardization consisted of "tagging" personal information such as addresses (tagged as placename), proper names (tagged as propername) and medication names (tagged as medicationname). Spelling errors where the intent of the participant could be



determined were corrected. If the error was beyond prediction that word was simply eliminated. All the responses were listed by group. Each corpus was then analyzed.

To determine whether native language (Spanish or English) and bilingual status impacts the selection of functionally relevant words a descriptive analysis of each corpus was made to include data for the top 50 most frequent words. The analysis included the word's frequency within the corpus, its lemma, its part of speech, its frequency in an external corpus (external corpus frequency), its concreteness rating, and its word length in phonemes. The minimum, maximum and mean external corpus frequency was determined. The frequency of each of the parts of speech was determined and entered into a table. Descriptive statistics of word length were also determined. Finally, the mean, minimum and maximum concreteness ratings were determined for the corpora with available concreteness ratings. The concreteness ratings for Spanish words were limited at the time of analysis, therefore the Spanish words were translated, and their concreteness was determined using the available English concreteness data. Differences between the corpora were documented.

To determine the variabilities which native language influences in lexicon selection, measurements for psycholinguistic differences between external corpus frequency, concreteness ratings, and word length in phonemes, were compared using inferential statistics. That data was compared across all language groups and in their question style (blank canvas versus open-ended questions). To further analyze the effects of native language on the selection of functionally relevant items, the top ten verbs and nouns of each language group within their question style category were extracted through the use of the corpus builder, and then were analyzed for corpus comparison.

To compare the words from this study's corpus to a larger world corpus of both the Spanish and English languages, word frequency distributions were compared to existing corpora: SUBTLEX and UBTLEX-ESP(Brysbaert New, 2009; Cuetos et al., 2011) and included within the inferential statistics analysis. Finally, to compare the part of speech occurrence of open ended questions vs. blank canvas questions, the part of speech frequency of each set was compared.

## Chapter 3

### Results

#### Participants

Following receipt of approval from the Institutional Review Board, the electronic survey was launched. Two hundred and sixty-seven participants completed the survey. The responses were sorted through by adhering to eligibility criteria. Three groups were created for analysis, a monolingual English group, a monolingual Spanish group, and a bilingual group. The final analyzed groups consisted of 81 Monolingual English speakers, 23 monolingual Spanish speakers, and 59 bilingual English/Spanish speakers. Participant demographics can be seen in Table. 1.

*Table 1: Participant Demographics*

Descriptor		English	Spanish	Bilingual
Total Number of Participants		81	23	59
Age	19-29	40	5	24
	30-39	9	5	20
	40-49	10	0	7
	50- 59	12	6	6
	60-69	8	7	2
	70-79	2	0	0
Native Language	English	81	0	16
	Spanish	0	23	43
Country of Origin	Canada	2		0
	Chile		1	5
	Colombia			4
	Ecuador			1
	United States	75		14
	Israel		1	1
	Honduras		19	22
	Mexico			3

	Philippines			1
	Puerto Rico			2
	Nicaragua	1		1
	Spain			1
	United Kingdom	1		1
	Venezuela		1	1
	Unspecified	2		2
Second Language	English	N/A	N/A	43
	Spanish	N/A	N/A	14
	Unspecified	N/A	N/A	2
Average Second Language Proficiency		N/A	N/A	4.23
Age of Acquisition	0-10	N/A	N/A	34
	11-20	N/A	N/A	20
	21-30	N/A	N/A	4
	31-40	N/A	N/A	1
	41-50	N/A	N/A	0
	51-60	N/A	N/A	0
Method of Acquisition	At Home Language	N/A	N/A	4
	Course	N/A	N/A	6
	School	N/A	N/A	42
	Other	N/A	N/A	7
Education Level	College Completed	34	8	24
	Doctorate Completed	10	1	3
	High School Completed	9	7	11
	Master	28	7	21
Gender	Female	66	16	51
	Male	14	6	8
	Other	1	1	0

## Descriptive Statistics

Data from the monolingual English group, monolingual Spanish group, and bilingual group were divided into 2 corpora each (a blank canvas corpus for each group and an open-ended question corpus for each group) for a total of six corpora: the Monolingual English Blank Canvas corpus (MEBC), and Monolingual English Open Ended Question corpus (MEOEQ), the Monolingual Spanish Blank Canvas corpus (MSBC) and Monolingual Spanish Open Ended Question corpus (MSOEQ), and the Bilingual Blank Canvas corpus (BBC) and Bilingual Open Ended Question corpus (BOEQ). The data cleaning and creation of each corpus was described within the methods section of this document.

To determine whether native language (Spanish or English) and bilingual status impacts the selection of functionally relevant words, a descriptive analysis of each corpus was made based on data for the top 50 most frequent words in each. These lists may be found in Appendix B. Table 2 displays the distributions of data within the open-ended question groups (MEOEQ, MSOEQ, BOEQ).

*Table 2: Open Ended Question Results*

<b>QUALIFIER</b>	<b>English Open-Ended Questions</b>	<b>Spanish Open-Ended Questions</b>	<b>Bilingual Open Ended Questions</b>
Total Tokens in Corpus	22661	5486	17199
Most Frequent Word	I	de	i
Most Used Language	English	Spanish	English
Most Frequent Part of Speech (MFPOS)	v	d and v	v
Appearance of (MFPOS)	14/50	9/50	11/50
Minimal External Corpus Frequency	1106.627451	140.1201923	1106.627451
Maximal External Corpus Frequency	41857.11765	33771.92308	41857.11765
Average External Corpus Frequency	9030.287443	8181.127404	12844.42265

Minimal Corpus Frequency Per Million	3000.750188	2734.232592	2674.574103
Maximal Corpus Frequency Per Million	61162.34941	44294.56799	43316.47189
Average Corpus Frequency Per Million	9506.200079	9019.32191	7749.287749
Minimal Concreteness Rating	1.35	1.35	1.35
Maximal Concreteness Rating	4.55	5	4.11
Average Concreteness Rating	2.259574468	2.4556	2.114242424
Most Frequent Word Length in Phoneme (MFWLIP)	2	2	2
Appearance of (MFWLIP)	25/50	17/50	30/50
Minimal Word Length in Phoneme	1	1	1
Maximal Word Length in Phoneme	4	7	4
Average Word Length in Phoneme	2.4	3.14	2.24

Note: The external corpus frequencies for the MEOEQ, BSOEQ were obtained from SUBTLEX, (Brysbaert New, 2009) . The frequency MSOEQ, was obtained from SUBTLEX-ESP (Cuetos et al., 2011). The concreteness ratings were all obtained from Brysbaert and New) 2014)

The table below, Table 3. Displays of the distributions of data within the blank canvas groups (MEBC, MSBC and BBC)

*Table 3: Blank Canvas Results*

<b>QUALIFIER</b>	<b>English Blank Canvas</b>	<b>Spanish Blank Canvas</b>	<b>Bilingual Blank Canvas</b>
Total Tokens in Corpus	1147.00	354.00	966.00
Most Frequent Word	you	no	"propername"
Most Used Language	English	Spanish	English
Most Frequent Part of Speech (MFPOS)	n	n	n
Appearance of (MFPOS)	14/50	27/50	35/50
Minimal External Corpus Frequency	12.29	17.96	19.33

Maximal External Corpus Frequency	41857.12	30664.78	41857.12
Average External Corpus Frequency	3350.43	2985.65	4209.53
Minimal Corpus Frequency Per Million	4359.20	5649.72	4140.79
Maximal Corpus Frequency Per Million	32258.06	25423.73	24844.72
Average Corpus Frequency Per Million	10479.51	10395.48	7619.05
Minimal Concreteness Rating	1.35	1.46	1.64
Maximal Concreteness Rating	5.00	5.00	5.59
Average Concreteness Rating	3.12	3.47	3.30
Most Frequent Word Length in Phoneme (MFWLIP)	3.00	4.00	4.00
Appearance of (MFWLIP)	17/50	12/50	12/50
Minimal Word Length in Phoneme	1.00	1.00	1.00
Maximal Word Length in Phoneme	6.00	8.00	7.00
Average Word Length in Phoneme	3.33	4.56	3.73

Note: The external corpus frequencies for the MEOEQ, BSOEQ were obtained from SUBTLEX, (Brysbaert New, 2009) . The frequency MSEOQ, was obtained from SUBTLEX-ESP (Cuetos et al., 2011). The concreteness ratings were all obtained from Brysbaert and New) 2014)

### Inferential Statistics

In addition to the descriptive analysis, a 2x3 multivariate analysis of variance was performed for the lists of fifty most frequent words in each corpus. The interaction of language group and corpus type was not significant ( $F(6, 556) = 0.992, p = 0.43$ ). The effect of corpus type was statistically significant ( $F(3, 278) = 44.297, p < .001$ ). The effect of language group also had a significant effect on the dependent variables ( $F(6, 556) = 6.749, p < .001$ ).

A test of between subject effects was also performed. There was a statistically significant effect of language group on external corpus frequency ( $F(2,280) = 3.071, p = 0.048$ ) and word length in phonemes ( $F(2,280) = 16.638, p < 0.001$ ), but not concreteness ( $F(2,280) = 1.779, p =$

0.171). There was also a significant effect of corpus type on external corpus frequency ( $F(1, 280) = 39.421, p < 0.001$ ), word length in phonemes ( $F(1, 280) = 73.522, p < 0.001$ ), and concreteness ( $F(1, 280) = 63.985, p < 0.001$ ).

A chi square test of independence was also run to examine the effect corpus type and language group on the ratio of function words to content words. Corpus type had a significant difference ( $\chi^2 = 98.895, p < 0.001$ ), but language group did not ( $\chi^2 = 2.755, p = 0.252$ ).

To further analyze the effects of native language on the selection of functionally relevant items, lists of the top ten verbs and nouns of each language group were also compared. These lists may be found in Appendix C. A 2x3 multivariate analysis of variance was performed on the top ten nouns according to language group and corpus type. The interaction of language group and corpus type was not significant. The effects of corpus type and language group were statistically significant (corpus type:  $F(3,48) = 3.248, p = .03$ ; language group:  $F(6, 96) = 2.901, p = .012$ ). A test of between subject effects was also performed. There was a statistically significant effect of corpus type on external corpus frequency ( $F(1, 50) = 4.103, p = .048$ ) and a statistically significant effect of language group on word length ( $F(1,50) = 8.443, p = 0.001$ ). Other between subject effects were not significant.

A 2x3 multivariate analysis of variance was performed on the top ten verbs according to language group and corpus type. The interaction of language group and corpus type was not significant. The effects of corpus type ( $F(3,52) = 3.737, p = 0.017$ ) and language group ( $F(6,104) = 5.208, p < 0.001$ ) were both statistically significant. A test of between subject effects was also performed. There was a statistically significant effect of corpus type on external corpus frequency ( $F(1,54) = 8.498, p = 0.005$ ) and concreteness ( $F(1,54) = 5.006, p = 0.029$ ). There



was a statistically significant effect of language group on external corpus frequency ( $F(2,54) = 5.540, p = 0.006$ ) and word length in phonemes ( $F(2,54) = 15.750, p < 0.001$ )

To compare the part of speech occurrence of open-ended questions vs. blank canvas questions a chi square test of independence was run to examine the effect corpus type and language group on the ratio of occurrence of part of speech. Corpus type had a significant difference ( $\chi^2 = 119.409, p < 0.001$ ), but language group did not ( $\chi^2 = 25.806, p = 0.057$ ).

## Chapter 4

### **Discussion**

The investigation describes the construction of six corpora: the Monolingual English Blank Canvas corpus (MEBC), the Monolingual English Open Ended Question corpus (MEOEQ), the Monolingual Spanish Blank Canvas corpus (MSBC), the Monolingual Spanish Open Ended Question corpus (MSOEQ), the Bilingual Blank Canvas corpus (BBC) and the Bilingual Open Ended Question Corpus (BOEQ), followed by analysis estimating the effects of the corpus groups and language groups on the top fifty words of each corpus and their psycholinguistic variables. It was hypothesized that language groups would significantly impact the variables of external corpus frequency, word length in phonemes, and concreteness. It was also hypothesized that corpus groups (blank canvas vs. OEQ) would impact concreteness and part of speech.

The initial descriptive analysis, which was carried out to determine whether native language (Spanish or English) and bilingual status impacts the selection of functionally relevant words, showed differences among language groups and differences among corpus groups. Notable differences included differences in word length in phonemes and external corpus frequencies. Furthermore, a difference of part of speech was observed in the MSOEQ. Unlike the MEOEQ corpus and the BOEQ corpus, the MSOEQ corpus displayed both verbs and determiners to be its most frequent part of speech. This difference of most frequent part of speech is consistent with the principle that subjects and objects in Romance languages such as Spanish are more likely to have determiners, whereas determiners are often more optional in English (Chierchia, 1998). It is observed that in Spanish sentence production determiners bring

immense context to noun phrases. The argument could be made that not training determiners during recovery could significantly decrease the coherence of the patient's dialogue.

The MANOVA results on the language groups effects on the corpora and their psycholinguistic variables showed that significant differences were observed between language groups and that language group impacted external corpus frequency and word length in phoneme. MEBC and MSBC displayed lower external frequencies than BBC. Similarly, BOEQ displayed the highest external corpus frequency where MEOEQ and MSOEQ displayed lower external frequencies. When considering a difference in native language and its impact on external corpus frequency it is understood that there is a significant difference caused by language. This difference in external corpus frequency denoted between the bilingual groups vs the monolingual groups could be due to the fact that the highly frequent words of each language combined increase the frequency of the overall dialogue, whereas monolinguals are able to cover a wider sample of frequencies. The translation of this difference into clinical practice and selection of functionally relevant items could suggest that when considering treatment options for a bilingual speaker, frequency ratings of target words may vary based on priority. The clinician and patient need to decide whether they must target the most frequent words of both languages or focus on a more ample variety of frequencies in one language alone.

MSBC group and MSOEQ displayed higher word length in phonemes than their bilingual and monolingual English speaker counterparts. It is possible that the nature of the Romance language, Spanish, with increased word length, would increase its phonemic complexity. Unlike English, Spanish is more phonetically spelled language (Delattre, 1945). Therefore, when a Spanish word is represented into graphemes the number of graphemes is generally equivalent to the number phonemes whereas in English the number of phonemes of a word tends to be smaller

than the number of graphemes. Research shows that “opaque languages promote a global reading strategy, and transparent languages force a local strategy” (De Leon Rodriguez, 2016, p. 7).

.Spanish is a more transparent language than English. When considering this fact we can further see why phonemic inventories may differ and how even developmental acquisition will differ between the languages. These differences could later influence recovery if a loss of language is present, and could also possibly influence functionally relevant item selection. Bilingual speakers have a double phonemic representation (Garcia-Sierra et al., 2012), so, the phonemic word length difference noted in Spanish and English would likely be accounted for in their dual phonemic inventory. When considering this difference in clinical practice a clinician must note that if the priority of recovery is to recover Spanish alone or both languages (English and Spanish) the functionally relevant items of Spanish will likely be more phonemically complex. A greater complexity could increase the difficulty of motor planning. This could lead to more failure in trials of set targets earlier in therapy but perhaps greater generalization of less phonemically complex items later.

When comparing the effects of corpus type and language group on the ratio of function words to content words. Corpus type had a significant difference, but language group did not. It was observed that the blank canvas corpus displayed a higher percentage of content words than the open-ended style questions. This finding is consistent with Renvall’s conclusion in 2013, suggesting that a blank canvas style of developing a list of functionally relevant items leads to more content words such as nouns. Therefore, in clinical practice, these results suggest that greater consideration must be given to the style of question that is used when selecting functionally relevant items.

The results from the comparison of the top ten verbs and nouns of each language group to further analyze the effects of native language on the selection of functionally relevant items displayed notable differences among the groups. When the top ten nouns were analyzed in each corpus group, it was determined that there was a statistically significant effect of language group on word length. The same basic results were also shown for verbs. Additionally, it was observed that there was a statistical significance of language group on the general frequency of verbs. Therefore, it is understood that native language impacts the complexity of phonemes and frequency of functionally relevant items.

Similarly, to the results observed in the inferential study of the groups as a whole, word length in phonemes appeared to be greater among the Spanish language group. When selecting functionally relevant items to rehabilitate language, complexity of phonemes must be considered, as greater complexity could increase the difficulty of motor planning. Therefore, when selecting content words (such as verbs and nouns) for a native Spanish speaker, one must determine whether the word length in phonemes is attainable for the subject attempting a recovery of motor planning. If the word length in phoneme is determined to be beyond what is achievable for a given client, a therapist may consider studying the topics that were assigned priority within the target list and decrease the complexity in phonemes of said target (through the use of synonyms) when possible, while maintaining the semantic content of the target.

Comparably, to the results observed in the inferential study of the groups as a whole, the bilingual groups displayed higher frequency rates. Once more this phenomenon may be due to the fact that bilingual speakers are dividing up their top words among two languages using the most frequent words of both languages instead of a broader distribution of one. It was observed that the blank canvas corpus displayed lower frequency ratings than the open-ended question

corpus. This may suggest that people do not necessarily pick the most frequent words as the most important words for their blank canvas selections.

Corpus type had a significant effect on the psycholinguistic variable of nouns and verbs. It was determined that there was a statistically significant effect of corpus type on external corpus frequency and concreteness for verbs and a statistically significant effect on external corpus frequency for nouns. The open-ended corpus verbs showed higher frequencies and lower concreteness ratings than their blank canvas counterparts. The blank canvas nouns showed lower frequency ratings. This may suggest that the blank canvas concreteness tendency stated by Renvall in 2013 mostly apply to verbs. As a reminder Renvall stated that blank canvas lists tend to have higher concreteness content words. However, in this study this effect is mainly seen among the verbs of the corpora. Perhaps the lower frequency ratings shown in the noun comparison results suggest that people are choosing nouns that are important to them within the blank canvas corpus, as opposed to more norm-typed frequent nouns. One may suggest that the nouns of the blank canvas corpora are more representative of open-ended usage. While verbs, tend to show higher concreteness ratings and frequencies, which follows the trends described by Renvall (2013).

The results from the comparison of part of speech occurrence between open ended questions vs. blank canvas questions showed that corpus type had a significant difference, but language group did not. This analysis showed that blank canvas corpora had the highest group of noun occurrence. On the other hand the results for open ended question corpus type showed that verbs were the most commonly appearing part of speech. These results coincide with Renvall statements regarding blank canvas part of speech tendencies (2013).

Throughout the analysis of the effects of corpus type on the psycholinguistic variables it was noted that corpus types had a significant effect of corpus type on external corpus frequency, concreteness and word length in phonemes. The open-ended corpus type exhibited higher external corpus frequency among the language groups. This effect may be due to the fact that the open-ended question corpora displayed determiners as a part of their most frequent part of speech which occur more often than content words in spoken language. When considering only content words (nouns and verbs) it was observed that corpus type had a significant effect on the external corpus frequency of verbs and nouns. Verbs appeared to be more frequent within the open-ended question groups. The previously stated effect of corpus style on all language groups and verbs may suggest that through the use of open-ended question style corpora for the selection of functionally relevant items one may find frequently used determiners and helping verbs to consider for inclusion within a target list. The concreteness effect caused by corpus type shows that blank canvas corpora displayed higher concreteness ratings. This effect was also seen on verbs across language groups. This effect is consistent with Renvall et al.'s observation (2013) that selection of functionally relevant items tends to display higher concreteness, limiting the expression of abstract thoughts for recovering patients (Bailey, 2016). The word length in phonemes effect caused by corpus type shows that blank canvas corpora displayed higher word lengths in phoneme. This effect was seen on nouns and verbs across language groups. This is an important consideration, given that a higher complexity of phonemes would lead to more difficulty acquiring target words. It is then that people are choosing more difficult words with higher concreteness. Although higher concreteness is usually associated with less difficulty, the fact remains that high concreteness continues to limit the communicative abilities of patients. Blank canvas leads to an unfortunate tradeoff between more phonemically complex low

concreteness words and phonemically simple high concreteness words. The overall heightened complexity of phonemes and concreteness rates seen in the blank canvas corpus type suggests that when people are forced to narrow down their communicative needs into a limited and decontextualized list. These findings are consistent with Renvall's conclusion in 2013. It is seen then that the blank canvas approach tends to limit the speakers' communication by not allowing their speech and language to fulfil their role as tools to achieve their goals as framed by the ICF.

Additional findings of the study suggest that there are certain differences in cultural values and communication priorities among the different language groups. For example a religious theme was observed more frequently within the Spanish corpora. In the MSBC, "dios" (god) and "iglesia" (church) appeared in the top 50 words, with "god" being the top 4<sup>th</sup> word within the corpus. On the other hand, "god" appeared as the 48<sup>th</sup> top word in the English corpus and 19<sup>th</sup> top word in the bilingual corpus. Moreover, the theme of family ranked highest among the Spanish corpora both in tokens and overall ranking. These differences may suggest different topical priorities among the groups. Although these particular differences are not indicative that all Spanish speaking populations would reflect such a variation, it does suggest that cultural differences may impact the psycholinguistic variables of the functionally relevant items selected for treatment.

When considering the implications of this study overall in clinical practice, there are a few factors to be considered. Native language will influence external frequency, and phoneme length of words. Corpus style (based on interview question strategy) will affect the concreteness, external frequency and phoneme length of words chosen. Regional differences may impact grammatical patterns regardless of the first language of the patient. Therefore as clinicians, we must consider the diversity and priorities of our clients to develop appropriately complex and



diverse target lists of functionally relevant items. It is important to also consider the manner of interviewing the patient: we must expand our question to be dialogue inducing and encompassing of the needs of the client. Open ended questions will be of essence to develop a list of diverse (in concreteness and phonemic complexity) verbs, but perhaps less necessary when selecting nouns. Selecting functionally relevant items will therefore consider client linguistic background and appropriate interviewing style.

### **Strengths and Limitations**

A strength of this study was the development of novel open-ended style questions which allowed for a dialogue heavy written reflection of participants' everyday conversations in a variety of realistic situations. This allowed for a more accurate study of the effect of language on the selection of functionally relevant items. Additionally, it allowed for further study of the effects of corpus type on the selection of functionally relevant items, validating statements that had been made but not verified with data (Bailey, 2016; Renvall, 2013a) that current practices of functionally relevant item selection limit the communicative needs of speakers. The questions used in this study could be later used to model how to obtain accurate samples of samples of communication for the development of functionally relevant item list.

It is acknowledged that primary limitations for this study were the differences in sizes across the different groups and the small sample sizes overall. The difference in sizes in groups affected the overall corpora sizes, which could impact the frequency of items in each corpus, hence impacting the comparative analysis. This study could be strengthened by larger and more heterogenous groups across all language groups. A second limitation to the study was the lack of available Spanish data regarding concreteness ratings. To compensate for this limitation Spanish words were translated and then imputed into the concreteness database used for the English

samples. Future research should include measuring concreteness ratings for more Spanish words to more accurately assign concreteness ratings to Spanish items in accordance with Spanish speaker perception.

## **Conclusion**

This study aimed to understand the effect of native language on functionally relevant item selection. The results of this study revealed that native language impacts the external corpus frequency and word length in phoneme of functionally relevant item. This difference, which is consistent with linguistic differences across language groups suggests that these differences should be considered when selecting functionally relevant items. Additionally, the findings of this study further support that blank canvas corpus types creates limitations in the selection of accurately representative functionally relevant items that could lead to better communication for speakers. As the world becomes increasingly diverse, accurate consideration of linguistic differences across speakers must be highlighted to ensure an inclusive, accurate and useful selection of functionally relevant items.

## References

- Albert, I., & Trommsdorff, G. (2014). The Role of Culture in Social Development Over the Lifespan: An Interpersonal Relations Approach. *Online Readings in Psychology and Culture*, 6(2). <https://doi.org/10.9707/2307-0919.1057>
- Ansaldò, A. I., Marcotte, K., Scherer, L., & Raboyeau, G. (2008). Language therapy and bilingual aphasia: Clinical implications of psycholinguistic and neuroimaging research. *Journal of Neurolinguistics*, 21(6), 539–557. <https://doi.org/10.1016/j.jneuroling.2008.02.001>
- Arimoto, T., Hiroaki, S., Masahiro, M., Hiromi, N., & Ryuichiro, H. (2019). Analysis of satisfaction and topics in repeated conversation through days. [http://semdial.org/anthology/Z19-Arimoto\\_semdial\\_0020.pdf](http://semdial.org/anthology/Z19-Arimoto_semdial_0020.pdf)
- Arbesman, S., Strogatz, S. H., & Vitevitch, M. S. (2010). Comparative Analysis of Networks of Phonologically Similar Words in English and Spanish. *Entropy*, 12(3), 327–337. <https://doi.org/10.3390/e12030327>
- Bailey, D. J. (2016). Development of an Aphasia Treatment Targeting Verbs with Low Concreteness. [Doctoral dissertation, The University of Utah].
- Balandin, S., & Iacono, T. (1998). A few well-chosen words. *Augmentative and Alternative Communication*, 14(3), 147–161. <https://doi.org/10.1080/07434619812331278326>
- Beukelman, D., McGinnis, J., & Morrow, D. (1991). Vocabulary selection in augmentative and alternative communication. *Augmentative and Alternative Communication*, 7(3), 171–185. <https://doi.org/10.1080/07434619112331275883>

- Boenisch, J., & Soto, G. (2015). The Oral Core Vocabulary of Typically Developing English-Speaking School-Aged Children: Implications for AAC Practice. *Augmentative and Alternative Communication, 31(1)*, 77–84.  
<https://doi.org/10.3109/07434618.2014.1001521>
- Bryen, D. N. (2008). Vocabulary to Support Socially-Valued Adult Roles. *Augmentative and Alternative Communication, 24(4)*, 294–301.  
<https://doi.org/10.1080/07434610802467354>
- Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41(4)* 977–990.
- Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46(3)*, 904–911.  
<https://doi.org/10.3758/s13428-013-0403-5>
- Charteris-Black, J., & Ennis, T. (2001). A comparative study of metaphor in Spanish and English financial reporting. *English for Specific Purposes, 20(3)*, 249–266.  
[https://doi.org/10.1016/S0889-4906\(00\)00009-0](https://doi.org/10.1016/S0889-4906(00)00009-0)
- Chierchia, G. Reference to Kinds across Language. *Natural Language Semantics 6*, 339–405 (1998). <https://doi.org/10.1023/A:1008324218506>
- Cuetos, F., Glez-Nosti, M., Barbón, A., & Brysbaert, M. (2011). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica, 32*, 133-143.

Delattre, P. (1945). Spanish Is a Phonetic Language. *Hispania*, 28(4), 511-516.

<https://doi.org/10.2307/333798>

De León Rodríguez, D., Buetler, K. A., Eggenberger, N., Laganaro, M., Nyffeler, T., Annoni, J., & Müri, R. M. (2016). The impact of language opacity and proficiency on reading strategies in bilinguals: An eye movement study. *Frontiers in Psychology*, 7.

<https://doi.org/10.3389/fpsyg.2016.00649>

Farrier, L., Yorkston, K., Marriner, N., & Beukelman, D. (1985). Conversational control in nonimpaired speakers using an augmentative communication system. *Augmentative and Alternative Communication*, 1(2), 65–73.

<https://doi.org/10.1080/07434618512331273551>

García-Sierra, A., Ramírez-Esparza, N., Silva-Pereyra, J., Siard, J., & Champlin, C. A. (2012). Assessing the double phonemic representation in bilingual speakers of Spanish and English: An electrophysiological study. *Brain and Language*, 121(3), 194–205.

<https://doi.org/10.1016/j.bandl.2012.03.008>

Gómez-Ruiz, M.I. (2010). Bilingüismo y cerebro: mito y realidad [Bilingualism and the brain: myth and reality]. *Neurologia*. 25(7):443-52. Spanish. Erratum in: *Neurologia*.

(2011)26(6)382. PMID: 20964991.

Goodglass, H., & Wingfield, A. (1997). Anomia: Neuroanatomical and cognitive correlates (H. Goodglass & A. Wingfield, Eds.; 1997-08827-000). *Academic Press*.

Imai, S., Walley, A. C., & Flege, J. E. (2005). Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and

- Spanish listeners. *The Journal of the Acoustical Society of America*, 117(2), 896–907.  
<https://doi.org/10.1121/1.1823291>
- Kim, H., Kintz, S., & Wright, H. H. (2020). Development of a measure of function word use in narrative discourse: Core lexicon analysis in aphasia. *International Journal of Language & Communication Disorders*, <https://doi.org/10.1111/1460-6984.12567>
- Lipski, J. (1986). The construction pa(ra) atrás among Spanish-English bilinguals: Parallel structures and universal patterns. *Iberoamericana* (1977-2000), 10(2/3 (28/29)), 87-96.  
Retrieved March 6, 2021, from <http://www.jstor.org/stable/41671040>
- Milman, L., Clendenen, D., & Vega-Mendoza, M. (2014). Production and integrated training of adjectives in three individuals with nonfluent aphasia. *Aphasiology*, 28(10), 1198–1222.  
<https://doi.org/10.1080/02687038.2014.910590>
- Ostrosky-Solis, F., Gutierrez, A. L., Flores, M. R., & Ardila, A. (2007). Same or different? Semantic verbal fluency across Spanish-speakers from different countries. *Arch Clin Neuropsychology*, 22(3), 367-377. <https://doi.org/10.1016/j.acn.2007.01.011>
- Palmer, R., Hughes, H., & Chater, T. (2017). What do people with aphasia want to be able to say? A content analysis of words identified as personally relevant by people with aphasia. *PLOS ONE*, 12(3), e0174065. <https://doi.org/10.1371/journal.pone.0174065>
- Pfaff, C. W. (1979). Constraints on Language Mixing: Intrasentential Code-Switching and Borrowing in Spanish/English. *Language*, 55(2), 291–318. JSTOR.  
<https://doi.org/10.2307/412586>

- Renvall, K., Nickels, L., & Davidson, B. (2013a). Functionally relevant items in the treatment of aphasia (part I): Challenges for current practice. *Aphasiology*, *27*(6), 636–650.  
<https://doi.org/10.1080/02687038.2013.786804>
- Renvall, K., Nickels, L., & Davidson, B. (2013b). Functionally relevant items in the treatment of aphasia (part II): Further perspectives and specific tools. *Aphasiology*, *27*(6), 651–677.  
<https://doi.org/10.1080/02687038.2013.796507>
- Raymer, A. M. (2005). Naming and word retrieval problems. In L. L. LaPointe (Ed.), *Aphasia and related neurogenic language disorders* (3rd ed., pp. 72–86). New York: Thieme.
- Sandberg, C., & Kiran, S. (2014). How justice can affect jury: Training abstract words promotes generalisation to concrete words in patients with aphasia. *Neuropsychological Rehabilitation*, *24*(5), 738–769. <https://doi.org/10.1080/09602011.2014.899504>
- Sandberg, C. W., & Gray, T. (2020). Abstract Semantic Associative Network Training: A replication and update of an abstract word retrieval therapy program. *American Journal of Speech-Language Pathology*, *29*(3), 1574–1595. [https://doi.org/10.1044/2020\\_AJSLP-19-00066](https://doi.org/10.1044/2020_AJSLP-19-00066)
- Stuart, S., Beukelman, D., & King, J. (1997). Vocabulary use during extended conversations by two cohorts of older adults. *Augmentative and Alternative Communication*, *13*(1), 40–47.  
<https://doi.org/10.1080/07434619712331277828>
- Thompson, C., & Shapiro, L. (2005). Treating agrammatic aphasia within a linguistic framework: Treatment of Underlying Forms. *Aphasiology*, *19*(10–11), 1021–1036.  
<https://doi.org/10.1080/02687030544000227>

- Vigliocco, G., Vinson, D. P., Druks, J., Barber, H., & Cappa, S. F. (2011). Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience & Biobehavioral Reviews*, *35*(3), 407–426. <https://doi-org.spot.lib.auburn.edu/10.1016/j.neubiorev.2010.04.007>
- Wisenburn, B., & Mahoney, K. (2009). A meta-analysis of word-finding treatments for aphasia. *Aphasiology*, *23*(11), 1338–1352. <https://doi.org/10.1080/02687030902732745>
- Worrall, L., Sherratt, S., Rogers, P., Howe, T., Hersh, D., Ferguson, A., & Davidson, B. (2011). What people with aphasia want: Their goals according to the ICF. *Aphasiology*, *25*(3), 309–322. <https://doi.org/10.1080/02687038.2010.508530>



## Appendix A: Survey Questions

### English Survey Questions

#### Blank Canvas Question

1. Imagine you have lost your ability to talk. If you could only use 25 words to communicate, what would they be?

#### Open-Ended Questions

##### Activities and Participation

1. Education: Imagine you are meeting someone for the first time. In a few sentences, write down what you would tell them about your educational past. (Please write your sentences in the way that you would most naturally talk to a stranger in an ongoing casual conversation.)
2. General Tasks and Demands: In a few sentences, describe what your normal weekday looks like.
3. General Tasks and Demands: You are at a doctor's appointment. The nurse asks you about your current health and medication intake. Please write a few sentences what you might say to the nurse.
4. Interpersonal relationships (self): Describe yourself in a few sentences (Please DO NOT include any personally identifying information such as name, address etc.)
5. Interpersonal relationships (friends): You are meeting a friend for coffee. In a few sentences, write the first things you might share to start the conversation. (Please write your

sentences in the way that you would most naturally talk to a friend in an ongoing casual conversation.)

7. Interpersonal relationships (strangers): You have just been introduced to a stranger. In a few sentences, write down what you say to introduce yourself and begin a conversation. (Please write your sentences as if you were addressing them to a stranger in an ongoing conversation.)

8. Interpersonal relationships (public figures): In a few sentences, describe how public figures influence your day-to-day life. (Please write your sentences in the way that you would most naturally talk to a acquaintance in an ongoing casual conversation.)

9. Interpersonal relationships (self): In a few sentences, write what you would like an acquaintance to know about your past. (Please write your sentences as if you were addressing them to an acquaintance in an ongoing conversation.)

10. Interpersonal relationships (self): In a few sentences, write what you would like to say about your future. (Please write your sentences in the way that you would most naturally talk to a acquaintance in an ongoing casual conversation.)

11. Social and integrated life in the community: In a few sentences, write down what you tell your friend about household routines. (Please write your sentences in the way that you would most naturally talk to a friend in an ongoing casual conversation.)

12. Social and integrated life in the community: You run into your friend and you engage in a conversation about games, sports and exercise. In a few sentences, write down things you might say during this conversation. (Please write your sentences in the way that you would most naturally talk to a friend in an ongoing casual conversation.)

13. Social and integrated life in the community: You run into your friend and you engage in a conversation about food. In a few sentences, please write down what you would say to your friend. (Please write your sentences in the way that you would most naturally talk to a friend in an ongoing casual conversation.)

14. Social and integrated life in the community: You run into your friend and you engage in a conversation about the weather. In a few sentences, please write down what you would say to your friend. (Please write your sentences in the way that you would most naturally talk to a friend in an ongoing casual conversation.)

15. Social and integrated life in the community: You run into your friend and engage in a conversation about public health (such as a pandemic). In a few sentences, please write down what you would say. (Please write your sentences in the way that you would most naturally talk to a friend in an ongoing casual conversation.)

16. Social and integrated life in the community: You run into your friend and you engage in a conversation about traveling. In a few sentences, please write down what you would say to them. (Please write your sentences in the way that you would most naturally talk to a friend in an ongoing casual conversation.)

17. Social and integrated life in the community: You run into your friend and you engage in a conversation about current news and events. In a few sentences, please write what you would say to your friend? (Please write your sentences in the way that you would most naturally talk to a friend in an ongoing casual conversation.)

18. Social and integrated life in the community: You run into your friend and you engage in a conversation about your hobbies. In a few sentences, please write down what you would say to

your friend. (Please write your sentences in the way that you would most naturally talk to a friend in an ongoing casual conversation.)

19. Social and integrated life in the community: You run into your friend and you engage in a conversation about your religion/church/worship. In a few sentences, please write down what you would say to your friend. (Please write your sentences in the way that you would most naturally talk to a friend in an ongoing casual conversation.)

20. Social and integrated life in the community: Someone is doing your weekly grocery shopping. In a few sentences, please write down what you would ask them to get for you? (Please write it as if you were talking to the person buying your groceries.)

21. Social and integrated life in the community: You run into your friend and you engage in a conversation about your morality, philosophy, and ethics. In a few sentences, please write down what you would say to your friend. (Please write your sentences in the way that you would most naturally talk to a friend in an ongoing casual conversation.)

22. Social and integrated life in the community: You run into your friend and you engage in a conversation about music. In a few sentences, please write down what you would say to your friend. (Please write your sentences in the way that you would most naturally talk to a friend in an ongoing casual conversation.)

23. Social and integrated life in the community: You run into your friend and you engage in a conversation about recreational activities. In a few sentences, please write down what you would say to them. (Please write your sentences in the way that you would most naturally talk to a friend in an ongoing casual conversation.)

24. Social and integrated life in the community: You run into your friend and you engage in a conversation about family life. In a few sentences, please write down what you would say to your friend. (Please write your sentences in the way that you would most naturally talk to your friend in an ongoing casual conversation.)

25. Social and integrated life in the community: You run into your friend and you engage in a conversation about work. In a few sentences write down what you would like your friend to know about your work life. (Please write your sentences in the way that you would most naturally talk to your friend in an ongoing casual conversation.)

26. Social and integrated life in the community: You run into your friend and you engage in a conversation about your emotional status. In a few sentences, please write down what you would say to your friend? (Please write your sentences in the way that you would most naturally talk to a friend in an ongoing casual conversation.)

27. Social and integrated life in the community: Someone has seriously offended you and hurt your feelings. Please write down what you would say to them (Please write your sentences in the way that you would most naturally respond to your offender in an ongoing conflict)

Environmental Factors:

28. Personal Consumption: You are telling you friend about all the technological devices you own. Please write down what you would say to your friend in a conversation about the technology you own. (Please write your sentences in the way that you would most naturally talk to a friend in an ongoing casual conversation.)

29. Civil Services: You have just encountered a person who is having a medical emergency. In a few sentences, please write down what you would say to the emergency operator in order to get help for the person in need.

30. Civil Services: Your neighbor's house is on fire. In a few words write down what you would say to the emergency operator.

31. Civil Services: You have just arrived at a new city in which you do not have a privately owned form of transportation. In a few sentences, please write what you would say to a local person to assist you in obtaining transportation

### **Spanish survey questions**

1. Educación: Imagine que se encuentra con alguien por primera vez. En unas pocas oraciones, escriba lo que les diría sobre su pasado educativo. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación)
2. Relaciones interpersonales: En unas pocas oraciones, describa cómo es su día normal de la semana.
3. Relaciones interpersonales (uno mismo): Descríbase a usted mismo en unas pocas frases.
4. Servicios civiles: La casa de su vecino está en llamas. En pocas palabras, escriba lo que le diría al operador de emergencia.
5. Consumo personal: Alguien está haciendo sus compras semanales. En unas pocas oraciones, escriba lo que les pediría que le dieran. Escríbalo como si estuviera hablando con la persona que compra sus alimentos.
6. Tareas y demandas generales: está en una cita con el médico. La enfermera le pregunta sobre su salud actual y la ingesta de medicamentos. Escriba algunas oraciones sobre lo que podría decirle a la enfermera.

7. Relaciones interpersonales (yo): En unas pocas oraciones, escribe lo que te gustaría decir sobre tu futuro. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación)
8. Servicios civiles: Acaba de llegar a una nueva ciudad en la que no tiene un medio de transporte privado. En unas pocas oraciones, escriba lo que le diría a una persona local para que lo ayude a obtener transporte.
9. Vida social e integrada en la comunidad: te encuentras con tu amigo y entablas una conversación sobre juegos, deportes y ejercicio. En unas pocas oraciones, escriba las cosas que podría decir durante esta conversación. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación)
10. Vida social e integrada en la comunidad: te encuentras con tu amigo y entablas una conversación sobre tu estado emocional. En unas pocas oraciones, escribe lo que le dirías a tu amigo. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación)
11. Servicios civiles: Acaba de encontrarse con una persona que tiene una emergencia médica. En pocas oraciones, escriba lo que le diría al operador de emergencia para obtener ayuda para la persona necesitada.
12. Consumo personal: le está contando a su amigo acerca de todos los dispositivos tecnológicos que posee. Escriba lo que le diría a su amigo en una conversación sobre la tecnología que posee. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación)
13. 13.Relaciones interpersonales (figuras públicas): en unas pocas frases, describe cómo las figuras públicas influyen en tu vida diaria. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación)

14. Vida social e integrada en la comunidad: te encuentras con tu amigo y entablas una conversación sobre el trabajo. En unas pocas frases, escribe lo que te gustaría que tu amigo supiera sobre tu vida laboral. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación)
15. Relaciones interpersonales (yo): En unas pocas oraciones, escribe lo que te gustaría que un conocido supiera sobre tu presente. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación)
16. Relaciones interpersonales (extraños): Le acaban de presentar a un extraño. En unas pocas oraciones, escriba lo que dice para presentarse y comenzar una conversación. (Escriba sus oraciones como si se las estuviera dirigiendo a un extraño en una conversación en curso).
17. Vida social e integrada en la comunidad: Te encuentras con tu amigo y entablas una conversación sobre la vida familiar. En unas pocas oraciones, escriba lo que le diría a su amigo. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación)
18. Relaciones interpersonales (yo): En unas pocas oraciones, escribe lo que te gustaría que un conocido supiera sobre tu pasado. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación)
19. Vida social e integrada en la comunidad: En unas pocas frases, escribe lo que le dices a tu amigo sobre las rutinas del hogar. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación)
20. Relaciones interpersonales (amigos): se encuentra con un amigo para tomar un café. En unas pocas oraciones, escribe las primeras cosas que podrías compartir para iniciar la conversación. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación).



21. Vida social e integrada en la comunidad: Te encuentras con tu amigo y entablas una conversación sobre noticias y eventos actuales. En unas pocas frases, escribe lo que le dirías a tu amigo. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación)
22. Vida social e integrada en la comunidad: Alguien lo ha ofendido gravemente y ha herido sus sentimientos. Por favor escriba lo que le diría (Por favor, escriba sus oraciones de la manera en que las diría en una conversación).
23. Vida social e integrada en la comunidad: te encuentras con tu amigo y entablas una conversación sobre música. En unas pocas oraciones, escriba lo que le diría a su amigo. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación).
24. Vida social e integrada en la comunidad: te encuentras con tu amigo y entablas una conversación sobre tus pasatiempos. En unas pocas oraciones, escribe lo que le dirías a tu amigo (Por favor, escriba sus oraciones de la manera en que las diría en una conversación)
25. Vida social e integrada en la comunidad: te encuentras con tu amigo y entablas una conversación sobre tu religión / iglesia / culto. En unas pocas oraciones, escriba lo que le diría a su amigo. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación).
26. Vida social e integrada en la comunidad: te encuentras con tu amigo y entablas una conversación sobre tu moralidad, filosofía y ética. En unas pocas oraciones, escriba lo que le diría a su amigo. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación).

27. Vida social e integrada en la comunidad: te encuentras con tu amigo y entablas una conversación sobre la comida. En unas pocas oraciones, escriba lo que le diría a su amigo. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación).
28. Vida social e integrada en la comunidad: te encuentras con tu amigo y entablas una conversación sobre el clima. En unas pocas oraciones, escriba lo que le diría a su amigo. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación).
29. Vida social e integrada en la comunidad: te encuentras con tu amigo y entablas una conversación sobre salud pública (como una pandemia). En unas pocas oraciones, escriba lo que diría. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación).
30. Vida social e integrada en la comunidad: te encuentras con tu amigo y entablas una conversación sobre viajes. En unas pocas oraciones, escriba lo que les diría. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación).
31. Vida social e integrada en la comunidad: te encuentras con tu amigo y entablas una conversación sobre actividades recreativas. En unas pocas oraciones, escriba lo que les diría. (Por favor, escriba sus oraciones de la manera en que las diría en una conversación).

### **Appendix B: Lists of Top 50 Most Frequent Words in Each Corpus**

#### *Appendix B 1: MEBC 50 Words*

<b>WORD</b>	<b>CORPUS FREQUENCY</b>	<b>LANGUAGE</b>	<b>PART OF SPEECH</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETE NESS RATING</b>	<b>WLIP</b>	<b>FUNCTION VS. CONTENT</b>
you	37	eng	pronoun	41857.12	4.11	2	function

i	28	eng	pronoun	39971.16	3.93	1	function
propername	24	eng	noun	//	//	//	function
no	23	eng	adverb	5971.55	2.45	2	function
help	22	eng	verb	921.12	2.56	4	content
me	19	eng	pronoun	9241.94	4.33	2	function
food	19	eng	noun	154.43	4.80	3	content
bathroom	17	eng	noun	61.67	4.52	6	content
yes	16	eng	adverb	1996.76	2.14	3	function
drink	15	eng	noun	247.39	4.76	5	content
help	15	eng	noun	921.12	2.56	4	content
love	24	eng	noun	1114.98	2.07	3	content
sad	14	eng	adjective	63.37	3.07	3	content
go	13	eng	verb	3793.04	3.15	2	content
want	12	eng	verb	2759.18	1.93	4	content
happy	12	eng	adjective	333.20	2.56	4	content
where	12	eng	adverb	1830.22	1.66	2	function
cold	12	eng	adjective	130.16	3.85	4	content
hungry	12	eng	adjective	77.08	2.90	5	content
please	11	eng	verb	1100.96	1.64	4	content
water	11	eng	noun	225.06	5.00	4	content
thank	11	eng	verb	1115.24	3.00	4	content
hot	11	eng	adjective	189.84	4.31	3	content
eat	10	eng	verb	251.88	4.44	2	content
good	10	eng	adjective	2610.14	1.64	3	content
home	10	eng	noun	774.33	4.11	3	content
pain	10	eng	noun	97.94	3.50	3	content
outside	9	eng	adverb	170.02	4.25	5	content
hurt	9	eng	verb	246.35	3.61	4	content
thirsty	9	eng	adjective	12.29	3.86	6	content
family	9	eng	noun	354.25	4.23	6	content
how	8	eng	adverb	3056.22	1.35	2	function

what	8	eng	pronoun	9842.45	2.00	3	function
yeah	8	eng	adverb	1996.76	2.14	3	function
sleep	8	eng	noun	227.94	4.44	4	content
when	8	eng	adverb	2034.10	1.60	3	function
stop	15	eng	verb	707.27	3.68	4	content
can	7	eng	verb	5247.45	4.55	3	function
tv	7	eng	noun	101.94	5.00	3	content
bad	7	eng	adjective	545.18	1.68	3	content
no	7	eng	adverb	5971.55	2.45	2	function
please	6	eng	adverb	1100.96	1.64	4	function
why	6	eng	adverb	2248.76	1.86	2	function
n't	6	eng	adverb	//	//	2	function
mom	6	eng	noun	430.39	4.40	3	content
god	6	eng	noun	903.16	2.07	3	content
hello	6	eng	interjection	585.43	2.80	4	content
do	6	eng	verb	6135.59	2.46	2	function
maybe	5	eng	adverb	926.45	1.52	4	function
sick	5	eng	adjective	165.43	2.97	3	content

*Appendix B 2: MEOEQ 50 Words*

<b>WORD</b>	<b>CORPUS FREQUENCY</b>	<b>LANGUAGE</b>	<b>PART OF SPEECH</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETE NESS RATING</b>	<b>WLIP</b>	<b>FUNCTION VS. CONTENT</b>
i	1386	eng	pronoun	39971.15686	3.93	1	function
and	695	eng	conjunction	13387.84314	1.52	3	function
the	657	eng	determiner	29449.17647	1.43	2	function
to	546	eng	preposition	22677.84314	1.55	2	function
a	517	eng	determiner	20415.27451	1.46	1	function
my	346	eng	pronoun	6762.72549	2.42	2	function
it	334	eng	pronoun	18896.31373	2.81	2	function
you	328	eng	pronoun	41857.11765	4.11	2	function
have	328	eng	verb	6161.411765	2.18	3	content

in	310	eng	preposition	9773.411765	3	2	function
is	304	eng	verb	9013	1.59	2	content
of	287	eng	preposition	11577.23529	1.67	2	function
do	234	eng	verb	6135.588235	2.46	2	content
n't	228	eng	adverb	//	//	2	function
for	217	eng	preposition	6895.098039	1.63	3	function
we	210	eng	pronoun	9011.901961	3.08	2	function
but	208	eng	conjunction	4417.470588	2.04	3	function
to	200	eng	preposition	22677.84314	1.55	2	function
are	185	eng	verb	5209.254902	1.96	2	content
'm	164	eng	verb	//	//	1	content
with	154	eng	preposition	5048.333333	2	3	function
so	153	eng	adverb	4244.156863	1.42	2	function
been	153	eng	verb	1736.72549	1.92	3	content
's	153	eng	verb	//	//	1	content
on	140	eng	preposition	6955.72549	3.25	2	function
get	137	eng	verb	4583.764706	2.38	3	function
am	133	eng	verb	1106.627451	1.96	2	content
at	132	eng	preposition	3217.098039	2.07	2	function
how	129	eng	adverb	3056.215686	1.35	2	function
be	112	eng	verb	5746.764706	1.85	2	function
what	111	eng	pronoun	9842.45098	2	3	function
really	107	eng	adverb	1500.156863	1.44	4	function
that	102	eng	preposition	14111.31373	1.54	3	function
me	101	eng	pronoun	9241.941176	4.33	2	function
or	99	eng	conjunction	1705.294118	1.72	2	function
this	95	eng	determiner	7978.72549	2.14	3	function
not	89	eng	adverb	5424.960784	2.08	3	function
they	89	eng	pronoun	4102.941176	2.93	2	function
that	87	eng	determiner	14111.31373	1.54	3	function
go	84	eng	verb	3793.039216	3.15	2	content

think	78	eng	verb	2691.392157	2.41	4	content
some	78	eng	determiner	1727.235294	2.48	3	function
was	76	eng	verb	5654.72549	1.69	3	function
can	76	eng	verb	5247.45098	4.55	3	function
just	73	eng	adverb	4749.137255	1.52	4	function
time	71	eng	noun	1958.627451	3.07	3	content
love	70	eng	verb	1114.980392	2.07	3	content
now	69	eng	adverb	3202.607843	1.48	2	function
good	68	eng	adjective	2610.137255	1.64	3	content
up	68	eng	preposition	3670	3.83		function

*Appendix B 3: MSBC 50 Words*

<b>WORD</b>	<b>CORPUS FREQUENCY</b>	<b>LANGUAGE</b>	<b>PART OF SPEECH</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETENESS RATING</b>	<b>WLIP</b>	<b>FUNCTION VS. CONTENT</b>
agua	9	esp	noun	249.93	5.00	4	content
no	9	esp	adverb	30664.78	2.45	2	function
nomb re	7	esp	noun	505.24	3.50	5	content
dios	7	esp	noun	1395.84	2.07	4	content
si	6	esp	adverb	5141.75	2.14	2	function
hola	6	esp	interjection	1960.46	2.80	4	content
baño	6	esp	noun	161.51	4.52	4	content
comi da	5	esp	noun	207.76	4.80	6	content
ayuda	5	esp	noun	321.78	2.56	5	content
dormi r	5	esp	verb	172.67	4.44	6	content
a	4	esp	preposition	23214.78	1.46	1	function
frío	4	esp	adjective	82.52	3.85	4	content
diner o	4	esp	noun	721.73	4.54	6	content
casa	4	esp	noun	1378.29	5.00	4	content

hambre	4	esp	noun	103.77	3.30	6	content
amor	4	esp	noun	394.76	2.07	4	content
sed	4	esp	noun	18.49	3.04	3	content
gracias	4	esp	noun	1961.30	2.15	7	content
familia	4	esp	noun	428.53	4.23	7	content
adiós	4	esp	noun	359.57	3.00	5	content
te	4	esp	pronoun	8240.19	4.11	2	function
necesito	4	esp	verb	703.49	1.69	8	content
mi	3	esp	pronoun	5761.01	2.42	2	function
tu	3	esp	pronoun	4002.69	4.11	2	function
triste	3	esp	adjective	82.12	3.07	6	content
dolor	3	esp	noun	123.49	3.50	5	content
calor	3	esp	noun	60.05	3.79	5	content
papá	3	esp	noun	754.13	4.29	4	content
mamá	3	esp	noun	787.00	4.40	4	content
me	3	esp	pronoun	9838.63	3.93	2	function
mañana	3	esp	adverb	637.91	3.44	6	content
bien	3	esp	adverb	5804.59	//	4	content
es	3	esp	verb	16779.42	1.59	2	function
vamos	3	esp	verb	2585.36	3.15	5	content
quiero	3	esp	verb	1973.49	1.93	6	content
amo	3	esp	verb	235.87	2.07	3	content
puedo	3	esp	verb	2024.06	4.55	5	content
un	2	esp	determiner	13081.08	3.97	2	function
permiso	2	esp	noun	99.64	2.27	7	content
trabajo	2	esp	noun	845.34	3.19	7	content
pipi	2	esp	noun	//	4.72	4	content
hijo	2	esp	noun	680.07	4.14	4	content
carro	2	esp	noun	17.96	4.89	5	content
música	2	esp	noun	140.12	4.31	6	content
biblia	2	esp	noun	22.19	4.61	6	content

iglesia	2	esp	noun	82.93	4.90	7	content
fuego	2	esp	noun	132.81	4.68	5	content
querer	2	esp	verb	50.55	1.93	6	content
comer	2	esp	verb	207.76	4.44	6	content
ver	2	esp	verb	1097.57	3.21	3	content

*Appendix B 4: MSOEQ 50 Words*

<b>WORD</b>	<b>CORPUS FREQUENCY</b>	<b>LANGUAGE</b>	<b>PART OF SPEECH</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETE NESS RATING</b>	<b>WLIP</b>	<b>FUNCTION VS. CONTENT</b>
de	243	esp	preposition	32894.69	1.84	2	function
y	186	esp	conjunction	16547.48	1.52	1	function
la	142	esp	determiner	21060.22	1.43	2	function
me	130	esp	pronoun	9838.63	3.93	2	function
en	119	esp	preposition	15578.00	3.25	2	function
a	115	esp	preposition	23214.78	1.55	1	function
que	98	esp	pronoun	33771.92	2.00	3	function
que	95	esp	conjunction	33771.92	2.00	3	function
mi	87	esp	determiner	5761.01	2.42	2	function
no	87	esp	adverb	30664.78	2.45	2	content
el	72	esp	determiner	17044.64	3.93	2	function
es	62	esp	verb	16779.42	1.59	2	function
los	59	esp	determiner	7082.02	1.43	3	function
las	54	esp	determiner	4633.39	1.43	3	function
para	50	esp	preposition	6524.21	1.63	4	function
con	47	esp	preposition	7173.49	2.00	3	function
por	43	esp	preposition	10498.97	1.63	3	function



una	42	esp	determi ner	8952.76	3.97	3	function
un	41	esp	determi ner	13081.08	3.97	2	function
gusta	39	esp	verb	600.10	1.89	5	content
se	37	esp	pronoun	7007.98	1.59	2	function
como	34	esp	conjunct ion	3595.84	1.35	4	function
mis	33	esp	determi ner	916.83	2.42	3	function
pero	32	esp	conjunct ion	5373.51	2.04	4	function
te	32	esp	pronoun	8240.19	4.11	2	function
lo	30	esp	determi ner	14864.57	1.43	2	function
soy	28	esp	verb	1879.50	1.96	3	function
hola	27	esp	interject ion	1960.46	2.80	4	function
muy	25	esp	adverb	2637.02	1.43	3	content
casa	24	esp	noun	1378.29	5.00	4	content
ya	22	esp	adverb	2667.43	1.48	2	content
tengo	22	esp	verb	2338.25	2.18	5	content
hay	22	esp	verb	2372.50	1.59	3	function
pues	21	esp	conjunct ion	339.66	1.44	4	function
si	21	esp	conjunct ion	5141.75	2.14	2	function
much o	20	esp	adverb	1239.62	1.69	5	content
al	19	esp	preposit ion	3731.39	1.55	2	function
del	19	esp	preposit ion	3747.31	1.84	3	function
estoy	19	esp	verb	2617.76	1.96	5	function
día	17	esp	noun	961.01	3.92	3	content
músi ca	17	esp	noun	140.12	4.31	6	content
nos	17	esp	pronoun	2137.96	3.08	3	function
tener	17	esp	verb	654.30	2.18	5	function
le	16	esp	pronoun	3339.04	4.11	2	function
bien	16	esp	adverb	5804.59	1.64	4	content
hacer	16	esp	verb	1827.81	2.46	5	function
trabaj o	15	esp	noun	845.34	3.19	7	content

hijos	15	esp	noun	211.73	4.89	5	content
yo	15	esp	pronoun	4513.53	3.93	2	function
ver	15	esp	verb	1097.57	3.21	3	content

*Appendix B 5: BBC 50 Words*

<b>WORD</b>	<b>CORPUS FREQUENCY</b>	<b>LANGUAGE</b>	<b>PART OF SPEECH</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETENESS RATING</b>	<b>WLIP</b>	<b>FUNCTION VS. CONTENT</b>
i	19	eng	pronoun	39971.16	3.93	1	function
propername	24	eng	noun	//	#N/A	//	function
no	20	eng/esp	adverb	5971.55	#N/A	2	function
you	18	eng	pronoun	41857.12	4.11	2	function
food	13	eng	noun	154.43	4.80	3	content
bathroom	12	eng	noun	61.67	4.52	6	content
love	10	eng	noun	1114.98	2.07	3	content
agua	10	esp	noun	249.93	5.00	4	content
hungry	9	eng	adjective	77.08	2.90	5	content
help	9	eng	noun	921.12	2.56	4	content
medication name	8	eng	noun	//	#N/A	//	content
water	8	eng	noun	225.06	5.00	4	content
baño	8	esp	noun	161.51	4.52	4	content
te	8	esp	pronoun	8240.19	3.93	2	content
quiero	8	esp	verb	1973.49	1.93	5	content
comida	8	esp	noun	207.76	4.80	6	content
yes	8	eng	adverb	1996.76	2.14	3	function
god	7	eng	noun	903.16	2.07	3	content
thank	7	eng	verb	1115.24	2.15	4	content
me	6	esp	pronoun	9838.63	4.33	2	function
pain	6	eng	noun	97.94	3.50	3	content
gracias	6	esp	interjection	1961.30	2.15	7	content

mamá	6	esp	noun	787.00	4.40	4	content
si	6	esp	adverb	5141.75	2.14	2	content
hola	6	esp	interjection	1960.46	2.80	4	content
want	6	eng	verb	2759.18	1.93	4	content
please	6	eng	adverb	1100.96	1.64	4	content
sleep	5	eng	noun	227.94	4.44	4	content
cold	5	eng	adjective	130.16	3.85	4	content
family	5	eng	noun	354.25	4.23	6	content
drink	5	eng	verb	247.39	4.76	5	content
hambre	5	esp	noun	103.77	3.30	5	content
familia	5	esp	noun	428.53	4.23	7	content
ayuda	5	esp	noun	321.78	2.56	5	content
amo	5	esp	verb	235.87	2.07	3	content
papá	5	esp	noun	754.13	4.29	4	content
gracia	5	esp	noun	1961.30	1.78	7	content
que	5	esp	pronoun	33771.92	2.00	2	content
do	5	eng	verb	6135.59	2.46	2	function
love	5	eng	verb	1114.98	2.07	3	content
what	5	eng	pronoun	9842.45	2.00	3	function
home	4	eng	adverb	774.33	4.11	3	content
hug	4	eng	verb	19.33	4.14	3	content
bye	4	eng	interjection	180.08	2.25	2	content
por	4	esp	preposition	10498.97	1.63	3	content
dormir	4	esp	verb	172.67	4.44	6	content
playa	4	esp	noun	57.64	4.79	5	content
go	4	eng	verb	3793.04	3.15	2	content
eat	4	eng	verb	251.88	4.44	2	content
where	4	eng	adverb	1830.22	1.66	2	function

*Appendix B 6: BOEQ 50 Words*

<b>WORD</b>	<b>CORPUS FREQUENCY</b>	<b>LANGUAGE</b>	<b>PART OF SPEECH</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETE NESS RATING</b>	<b>WLIP</b>	<b>FUNCTION VS. CONTENT</b>
i	745	eng	pronoun	39971.16	3.93	1	function

the	422	eng	determiner	29449.18	1.43	2	function
a	385	eng	determiner	20415.27	1.46	1	function
and	361	eng	conjunction	13387.84	1.52	3	function
to	292	eng	preposition	22677.84	1.55	2	function
que	245	esp	pronoun	33771.92	2.00	2	function
y	197	esp	conjunction	16547.48	1.52	1	function
it	196	eng	pronoun	18896.31	2.81	2	function
de	194	esp	preposition	32894.69	1.84	2	function
my	184	eng	pronoun	6762.73	2.42	2	function
you	182	eng	pronoun	41857.12	4.11	2	function
me	170	eng/esp	pronoun	9838.63	2.97	2	function
in	153	eng	preposition	9773.41	3.00	2	function
is	145	eng	verb	9013.00	1.59	2	function
do	134	eng	verb	6135.59	2.46	2	function
of	130	eng	preposition	11577.24	1.67	2	function
have	128	eng	verb	6161.41	2.18	3	function
la	127	esp	determiner	21060.22	1.43	2	function
'm	107	eng	verb	//	//	1	function
's	107	eng	verb	//	//	1	function
en	104	esp	preposition	15578.00	3.25	2	function
but	95	eng	conjunction	4417.47	2.04	3	function
for	94	eng	preposition	6895.10	1.63	3	function
el	93	esp	determiner	17044.64	3.93	2	function
mi	91	esp	pronoun	5761.01	#N/A	2	function
no	90	eng/esp	adverb	30664.78	2.45	2	function
n't	86	eng	adverb	//	//	2	function
with	80	eng	conjunction	5048.33	2.00	3	function
we	79	eng	pronoun	9011.90	3.08	2	function
are	77	eng	verb	5209.25	1.96	2	function
what	75	eng	pronoun	9842.45	2.00	3	function

that	74	eng	determi ner	14111.31	1.54	3	function
so	68	eng	adverb	4244.16	1.42	2	function
this	67	eng	determi ner	7978.73	2.14	3	function
get	63	eng	verb	4583.76	2.38	3	content
on	62	eng	preposit ion	6955.73	3.25	2	function
una	61	esp	determi ner	8952.76	3.97	3	function
at	58	eng	preposit ion	3217.10	2.07	2	function
am	58	eng	verb	1106.63	1.96	2	function
un	58	esp	determi ner	13081.08	3.97	2	function
not	57	eng	adverb	5424.96	2.08	3	function
para	57	esp	preposit ion	6524.21	1.63	4	function
lo	57	esp	determi ner	14864.57	1.43	2	function
was	54	eng	verb	5654.73	1.69	2	function
about	54	eng	preposit ion	3631.49	1.77	4	function
how	53	eng	adverb	3056.22	1.35	2	function
be	52	eng	verb	5746.76	1.85	2	function
like	50	eng	verb	3998.96	1.89	3	content
es	47	esp	verb	16779.42	1.59	2	function
that	46	eng	determi ner	14111.31	1.54	3	function

## Appendix C: Lists of Top 10 Most Frequent Nouns and Verbs in Each Corpus

### Appendix C 1: MEBC Nouns

<b>CORPUS TYPE</b>	<b>LANGUAGE GROUP</b>	<b>WORD</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETENESS RATINGS</b>	<b>WLI P</b>
bc	eng	propersname	N/A	N/A	N/A
bc	eng	food	154.43	4.80	3
bc	eng	bathroom	61.67	4.52	6
bc	eng	drink	247.39	4.76	5
bc	eng	help	921.12	2.56	4
bc	eng	love	1114.98	2.07	3
bc	eng	water	225.06	5.00	5
bc	eng	pain	97.94	3.50	3
bc	eng	home	774.33	4.11	3
bc	eng	family	354.25	4.23	6

### Appendix C 2: MEOEQ Nouns

<b>CORPUS TYPE</b>	<b>LANGUAGE GROUP</b>	<b>WORD</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETENESS RATINGS</b>	<b>WLI P</b>
oe	eng	time	1958.63	3.07	3
oe	eng	people	1102.98	4.82	5
oe	eng	day	801.82	3.92	2
oe	eng	school	333.12	4.79	4
oe	eng	family	354.25	4.23	6
oe	eng	work	798.02	3.48	4
oe	eng	home	774.33	4.11	3
oe	eng	house	514.00	5.00	3
oe	eng	things	692.88	N/A	4
oe	eng	life	796.65	2.69	3

### Appendix C 3: MSBC Nouns

<b>CORPUS TYPE</b>	<b>LANGUAGE GROUP</b>	<b>WORD</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETENESS RATINGS</b>	<b>WLI P</b>
bc	esp	agua	249.93	5.00	4
bc	esp	dios	1395.84	2.07	4

bc	esp	nombre	505.24	3.50	6
bc	esp	baño	161.51	4.52	4
bc	esp	ayuda	321.78	2.56	5
bc	esp	comida	207.76	4.80	6
bc	esp	dinero	721.73	4.54	6
bc	esp	gracias	1961.30	2.15	7
bc	esp	familia	428.53	4.23	7
bc	esp	amor	394.76	2.07	4

*Appendix C 4: MSOEQ Nouns*

<b>CORPUS TYPE</b>	<b>LANGUAGE GROUP</b>	<b>WORD</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETENESS RATINGS</b>	<b>WLI P</b>
oe	esp	casa	1378.29	5.00	4
oe	esp	musica	140.12	4.31	6
oe	esp	dia	961.01	3.92	3
oe	esp	hijos	211.73	4.89	5
oe	esp	tiempo	1296.97	3.07	6
oe	esp	trabajo	845.34	3.48	7
oe	esp	personas	288.25	4.82	8
oe	esp	familia	428.53	4.23	7
oe	esp	direccion	93.41	3.89	8
oe	esp	vida	1104.38	2.69	4

*Appendix C 5: BBC Nouns*

<b>CORPUS TYPE</b>	<b>LANGUAGE GROUP</b>	<b>WORD</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETENESS RATINGS</b>	<b>WLI P</b>
bc	biling	propername	N/A	N/A	N/A
bc	biling	food	154.43	4.80	3
bc	biling	bathroom	61.67	4.52	6
bc	biling	love	1114.98	2.07	3
bc	biling	agua	249.93	5.00	4
bc	biling	help	321.78	2.56	4
bc	biling	water	225.06	5.00	5
bc	biling	quiero	1973.49	1.93	5
bc	biling	medicationame	N/A	N/A	N/A
bc	biling	comida	207.76	4.80	6

*Appendix C 6: BOEQ Nouns*

<b>CORPUS TYPE</b>	<b>LANGUAGE GROUP</b>	<b>WORD</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETENESS RATINGS</b>	<b>WLI P</b>
oe	biling	time	1958.63	3.07	3
oe	biling	family	354.25	4.23	6
oe	biling	casa	1378.29	5.00	4
oe	biling	home	774.33	4.11	3
oe	biling	day	801.82	3.92	2
oe	biling	life	796.65	2.69	3
oe	biling	vida	1104.38	2.69	4
oe	biling	house	514.00	5.00	3
oe	biling	people	1102.98	4.82	5
oe	biling	work	798.02	3.48	4

*Appendix C 7: MEBC Verbs*

<b>CORPUS TYPE</b>	<b>LANGUAGE GROUP</b>	<b>WORD</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETENESS RATINGS</b>	<b>WLI P</b>
bc	eng	help	921.12	2.56	4
bc	eng	love	1114.98	2.07	3
bc	eng	thank	1115.24	3.00	4
bc	eng	go	3793.04	3.15	2
bc	eng	want	2759.18	1.93	4
bc	eng	please	1100.96	1.64	4
bc	eng	eat	251.88	4.44	2
bc	eng	hurt	246.35	3.61	4
bc	eng	stop	707.27	3.68	4
bc	eng	need	1294.90	1.69	3

*Appendix C 8: MEOEQ Verbs*

<b>CORPUS TYPE</b>	<b>LANGUAGE GROUP</b>	<b>WORD</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETENESS RATINGS</b>	<b>WLI P</b>
oe	eng	get	4583.76	2.38	3
oe	eng	go	3793.04	3.15	2



oe	eng	think	2691.39	2.41	4
oe	eng	love	1114.98	2.07	3
oe	eng	going	2123.29	2.69	4
oe	eng	work	798.02	3.48	4
oe	eng	know	5721.18	1.68	2
oe	eng	like	3998.96	1.89	3
oe	eng	doing	1029.25	2.66	4
oe	eng	feel	627.24	2.28	3

*Appendix C 9: MSBC Verbs*

<b>CORPUS TYPE</b>	<b>LANGUAGE GROUP</b>	<b>WORD</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETENESS RATINGS</b>	<b>WLI P</b>
bc	esp	dormir	172.67	4.44	6
bc	esp	puedo	2024.06	4.55	5
bc	esp	quiero	1973.49	1.93	5
bc	esp	necesito	703.49	1.69	8
bc	esp	comer	196.73	4.44	5
bc	esp	amo	235.87	2.07	3
bc	esp	vamos	2585.36	3.15	5
bc	esp	duele	71.51	3.61	5
bc	esp	tener	654.30	2.18	5
bc	esp	salir	413.92	3.93	5

*Appendix C 10: MSBOEQ Verbs*

<b>CORPUS TYPE</b>	<b>LANGUAGE GROUP</b>	<b>WORD</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETENESS RATINGS</b>	<b>WLI P</b>
oe	esp	gusta	600.10	1.89	5
oe	esp	tengo	2338.25	2.18	5
oe	esp	ver	1097.57	3.21	3
oe	esp	puedo	2024.06	4.55	5
oe	esp	gustar	249.23	1.89	8
oe	esp	ayudar	166.63	2.56	6
oe	esp	veo	388.58	3.21	3
oe	esp	creo	1715.34	1.55	4
oe	esp	tomar	255.48	3.06	5
oe	esp	ir	980.41	3.15	2

*Appendix C 11: BBC Verbs*

<b>CORPUS TYPE</b>	<b>LANGUAGE GROUP</b>	<b>WORD</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETENESS RATINGS</b>	<b>WLI P</b>
bc	biling	want	2759.18	1.93	4
bc	biling	thank	1115.24	3.00	4
bc	biling	love	1114.98	2.07	3
bc	biling	please	1100.96	1.64	4
bc	biling	go	3793.04	3.15	2
bc	biling	hurt	246.35	3.61	4
bc	biling	eat	251.88	4.44	2
bc	biling	sleep	227.94	4.44	4
bc	biling	shower	41.12	4.89	4
bc	biling	like	3998.96	1.89	3

*Appendix C 12: BOEQ Verbs*

<b>CORPUS TYPE</b>	<b>LANGUAGE GROUP</b>	<b>WORD</b>	<b>EXTERNAL CORPUS FREQUENCY</b>	<b>CONCRETENESS RATINGS</b>	<b>WLI P</b>
oe	biling	get	4583.76	2.38	3
oe	biling	like	3998.96	1.89	3
oe	biling	going	2123.29	2.69	4
oe	biling	know	5721.18	1.68	2
oe	biling	go	3793.04	3.15	2
oe	biling	take	1891.04	3.06	3
oe	biling	got	3306.49	1.93	3
oe	biling	love	1114.98	2.07	3
oe	biling	think	2691.39	2.41	4
oe	biling	need	1294.90	1.69	3