The Effects of Frequency on Dual-Route Versus Single-Route Processing of Morphologically Complex Terms: A Usage-Based Experiment

Guinevere J. Deaver
Brigham Young University - Provo

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

Part of the Spanish and Portuguese Language and Literature Commons

BYU ScholarsArchive Citation
Deaver, Guinevere J., "The Effects of Frequency on Dual-Route Versus Single-Route Processing of Morphologically Complex Terms: A Usage-Based Experiment" (2013). All Theses and Dissertations. 4282.
https://scholarsarchive.byu.edu/etd/4282

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen amatangelo@byu.edu.
The Effects of Frequency on Dual-Route Versus Single-Route
Processing of Morphologically Complex Terms:
A Usage-Based Experiment

Guinevere Hand Deaver

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of
Masters of Arts

Jeffrey S. Turley, Chair
Orlando Alba
David Eddington

Department of Spanish and Portuguese
Brigham Young University
December 2013

Copyright © 2013 Guinevere Hand Deaver
All Rights Reserved
With the availability of frequency dictionaries, such as Alameda and Cuetos (1995) or the *Corpus del Español* (2002), it is now possible to explore the effects of frequency on linguistic items. The following is a study exploring the effects of frequency on Spanish affixed words. While the debate of dual-route versus single-route processing continues, the results of this study suggest that L2 Spanish speakers use a dual-route model and decompose morphologically complex words when the base frequency is higher than the surface frequency. L2 Spanish speakers perceive derived words with a higher base frequency as more complex than derived words with a lower base frequency. The results of this study do not suggest the same process occurs for native Spanish speakers. When asked to identify the more complex word of a pair, native Spanish speakers are just as likely to select the derived word with a lower base frequency as they are to select the derived word with a higher base frequency suggesting a single-route model.
ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to all who have assisted me in this endeavor. Many thanks to my thesis advisor, Dr. Turley, for his counsel and encouragement throughout this process. Also, thanks to my committee members for their guidance and support. I thank Dr. Eddington for sharing his expertise and excitement for learning. I also thank Dr. Alba who provided a perspective that proved very valuable. I give thanks to the members of the Department of Spanish and Portuguese for their continued encouragement and support as I have pursued this degree.

I express my sincere thanks for the participants from the ELC and the 300- and 400-level Spanish classes who took the survey. Without their participation, this study could not have been completed. I also thank the native Spanish-speaker graduate students of the Department of Spanish and Portuguese for taking the pilot survey so that the final survey could be solidified.

And finally, I express my gratitude to my husband, parents, and numerous friends who endured this process with me and were always ready to encourage me and who offered support and love. I thank my family for believing I could accomplish such an undertaking. And thanks to my father, for always challenging me to accomplish more, and my mother, for always helping me work out my thoughts. I also express my gratitude to my husband who has taken care of so much so that I could focus on my research. I am truly blessed and grateful.
TABLE OF CONTENTS

Chapter 1 – Introduction ................................................................................................................. 1
  Statement of the problem ............................................................................................................ 1
  Justification of the problem ....................................................................................................... 2
  Delimitation of the problem ........................................................................................................ 3

Chapter 2 – Review of the Literature .............................................................................................. 4
  Usage-based Theoretical Framework .......................................................................................... 4
  Frequency .................................................................................................................................... 4
  Dual Route versus Single Route .................................................................................................. 5

Chapter 3 – Methodology ............................................................................................................... 7
  Part 1 – Word Selection .............................................................................................................. 8
    Prefixed Word Selection .......................................................................................................... 9
    Suffixed Word Selection ....................................................................................................... 12
    Pairing of Derived Words ...................................................................................................... 13
  Part 2 – Survey of Spanish Speakers ......................................................................................... 16
    Pilot Survey ........................................................................................................................... 16
    Final Survey ........................................................................................................................... 23
    Participants ............................................................................................................................ 25

Chapter 4 – Results ....................................................................................................................... 27
  Results from the survey ............................................................................................................. 27
  Discussion ......................................................................................................................................... 29
  Semantic Transparency Analysis .............................................................................................. 31
LIST OF TABLES

Table 1  Frequency of Selected Prefixes ................................................................. 10
Table 2  Frequency of High-frequency Suffixes ....................................................... 12
Table 3  Frequency of Low-frequency Suffixes ......................................................... 13
Table 4  Finalized Prefixed Pairs in the Survey ......................................................... 21
Table 5  Finalized Suffixed Pairs in the Survey ......................................................... 21
Table 6  Finalized Filler Pairs in the Survey ............................................................. 22
Table 7  Survey Results of Wilcoxon Test ................................................................. 29
LIST OF FIGURES

Figure 1. Sample Spreadsheet of the Prefix des-.............................................................. 11

Figure 2. The Prefix sub-.................................................................................................. 15

Figure 3. The Suffix -dor.................................................................................................. 15

Figure 4. View of Excel Spreadsheet with Word A, the Derived Form with a Higher Derived Frequency, Compared to Word B, the Derived Form with a Higher Base Frequency ..... 18

Figure 5. Graph Depicting the Log of the Ratio of Derived Frequencies of Prefixed Pairs .... 19

Figure 6. Graph of Log of Ratio of Derived Frequencies of 15 Prefixed Pairs for Survey .... 19

Figure 7. Graph Depicting the Log of the Ratio of Derived Frequencies of Suffixed Pairs .... 20

Figure 8. Graph of Log of Ratio of Derived Frequencies of 15 Suffixed Pairs for Survey .... 20

Figure 9. Survey Instructions .................................................................................................. 24
Chapter 1 – Introduction

This study investigates the role of frequency in the decomposition of derived Spanish words. Derived words contain a base, which consists of a root or a root with an affix, and one or more affixes (Hualde, Olarrea, & Escobar, 2001). An example of a derived word is papelera, which contains the base or root papel and the affix -era. Nacionalismo is also an example of a derived word. The suffix -al is added to the base nación-, which is also the root, and the suffix – ismo is added to the base nacional-. The present study provides evidence as to how the frequency of the base relates to the processing of the derived word. The question is, when a person hears or reads a derived word, do they break the derived word into the base and the affix, or do they process it as a whole word?

Statement of the problem

With the increased availability of corpora, or collections of texts, it has become increasingly plausible to conduct linguistic research with a usage-based approach. The usage-based model looks at actual language use for patterns and connections that are directly related to frequency. Jennifer Hay (2001) investigates the effects of frequency on morphologically complex English terms and morphological decomposition. She argues that “it is relative frequency [the frequency relationship between tokens in the lexicon], rather than absolute frequency [the actual frequency of a token in the lexicon], that affects the decomposability of morphologically complex words” (Hay, 2001, p.1041). The results of her experiment show that “subjects perceive derived forms that are more frequent than their bases to be significantly less complex than matched counterparts that are less frequent than their bases” (p. 1041). She also shows that “a low-frequency form is likely to be nontransparent if it is composed of even-lower-frequency parts. And a high-frequency form may be highly decomposable if the base word it
contains is higher frequency still” (p. 1041). Hay uses the dictionary to reveal that semantic drift was more likely in derived forms that have a higher frequency than their bases. Derived forms with a higher frequency base are less prone to semantic drift (Hay, 2001).

Hay’s experiment examines frequency and derivational morphology in English. Do the same trends exist in Spanish? The purpose of the present study is to investigate the effects of frequency on morphologically complex terms of Spanish in order to determine how frequency affects morphological processing. I ask the same question that Hay (2001) asked: “Do we decompose affixed words upon encountering them, breaking them down into their parts in order to access lexical entries associated with their component morphemes? Or do we access affixed words as wholes, accessing an independent lexical entry?” (p.1044). Are affixed Spanish words processed through a dual-route model (that is, a model where some complex words are decomposed and others are not), or is the whole-word representation processed through a single-route model in which all words are processed the same way?

Justification of the problem

Many arguments have been presented in recent years in the discussion of dual-route and single-route models of processing. While I have been able to find research examining English with the Usage-Based Model, I have not found as much with Spanish, especially morphological studies. There is also a paucity of studies on how speakers of Spanish store and process language. I believe that relative frequency affects processing and that people use different processes based on frequency. In the research I have performed thus far, I have discovered new insights into my own thinking and language processing. I believe valuable information is to be learned through such an experience.
**Delimitation of the problem**

While some research has been conducted on Spanish inflectional morphology, I will focus my research on derivational morphology. While inflectional morphology is a factor in this study, it will only be considered in the selection of terms. Only singular terms will be used. In addition, the masculine form of adjectives will be included in the lists and survey as it is unmarked for gender. Finally, infinitives will form part of the study, but conjugations will not. Though it would be interesting to research compounding as well, the scope of such a task would be too broad. While many affixes will be included in this experiment, others will not. I will not be researching prefixoids such as *tele-*, *radio-* or *video-* nor emotive suffixes such as *-ito*, *-ete*, *-azo*, and so on. Prefixoids generate scientific or technical vocabulary and their status as prefixes is controversial (Lang, 1990).

In conclusion, the purpose of this experiment is to provide insight into affixed forms, the role of frequency, and processing routes.
Chapter 2 – Review of the Literature

Usage-based Theoretical Framework

Usage-based linguistics studies the role of actual language usage. Three key concepts of usage-based linguistics, as outlined by Tummers, Heylen, and Geeraerts (2005), include “the priority assigned to language use, the integration of competence and performance, and the rejection of the rule-list fallacy” (p. 228). In addition to the importance of language use, “the emergence of grammar from usage data presupposes an interaction with other cognitive capacities, such as abstraction, perception, and learning” (Tummers et al., 2005, p. 228). Bybee (2001) states that “the brain operates in the same way in different domains” (p. 7). And according to Tomasello (2003), “psychologists and cognitive scientists no longer think of children’s learning as isolated association-making and induction, but rather they think of it as integrated with other cognitive and social-cognitive skills—in ways that Skinner and the Behaviorist (and Chomsky in his critiques) could never have envisaged” (p. 3). The patterns, mental representations, and competence of language come from language use and the employment of cognitive abilities. Tomasello (2003) adds that “language structure emerges from language use” (p. 5).

Frequency

Frequency plays an important role in usage-based linguistics as it deals with language usage. According to Tummers et al. (2005), “frequency counts of different sorts are probably the most wide-spread quantification technique in usage-based models of language” (p. 240). There are many ways to count frequency (Baayen & Schreuder, 2004). Bybee (2001) focuses on two ways to count frequency, token and type frequency. Token frequency is the number of times a word occurs in a language sample. Type frequency is the number of times a pattern occurs in a
language sample. Bybee (2001) discusses the role of frequency in a usage-based model. She states that “high-frequency words and phrases have stronger representations in the sense that they are more easily accessed and less likely to undergo analogical changes” and that “low-frequency words are more difficult to access” (p. 6). Hay (2001) comments that high frequency forms tend not to be decomposed, but accessed whole. Hay also states that current models of morphological processing “predict that derived forms that are more frequent than their bases should be less decomposable than derived forms that are less frequent than their base, regardless of the absolute frequency of the derived form” (Hay, 2001, p. 1042).

In contrast to previous theories using rules or lists to describe the lexicon, a usage-based theory predicts that words or patterns are stored in the lexicon though connections. The goal of a usage-based approach is to discover or uncover generalizations regarding linguistic units, described as schemas, which are used to categorize and store linguistics items (Bybee, 2001). The lexicon involves a network of connections. Thus, derived forms and their bases can be connected through the base that is common to both. Nación, nacional, and nacionalismo are all connected through nación. With connections such as these, how does decomposition occur in conjunction with frequency? Hay (2001) states “while it is widely assumed that high-frequency morphologically complex forms tend to display characteristics of noncompositionality, models of morphological processing do not predict a direct relationship between absolute frequency and decomposition. Rather, they predict a relationship between decomposition and the relative frequency of the derived form and the base” (p. 1041).

**Dual Route versus Single Route**

Morphological decomposition has been the topic of the ongoing debate regarding dual-route versus single-route processes. The main premise of the dual-route model is that regular
morphological words are processed by rules of concatenation, the rules for linking word parts together, and that irregular morphological words are processed as whole words. On the other hand, the main premise of the single-route model is that all words are processed the same way regardless of having regular or irregular morphology. Eddington (2009) states that the controversy between dual-route and single-route processing originally centered on the processing of the English past tense because it displays both regular inflection patterns, such as wash- washed, and irregular inflectional patterns, such as sing-sang. Children tend to extend the -ed ending to irregularly inflected verbs, suggesting that the irregular form must be stored in memory while regular forms follow a suffixation rule (Pinker and Ullman, 2002). While the English past tense has been a topic of interest, Eddington (2009) notes that “morphological processing in other languages such as German and Italian have been cited as evidence for and against both models as well” (p. 174). Eddington (2009) reviewed four studies that suggest a dual-route model for Spanish verbal inflection and demonstrated that the single-route model could work as well. Thus, the debate over dual-route versus single-route continues.
Chapter 3 – Methodology

In regards to the processing of affixed Spanish words, I predict that the derived forms, whose whole-word frequency is higher than the base frequency, will be processed as whole words and the derived forms whose base frequency is higher than the whole-word frequency will be decomposed. In order to test this hypothesis, a study involving two parts will be conducted.

The first part of this study involves a usage-based approach to morphology using corpora for the selection of derived word forms based on frequency. A list of derived forms will be compiled, together with their whole-word frequencies as well as their base frequencies. The frequencies will be culled from the Alameda and Cuetos (1995) frequency dictionary. The bases will come from information found in the Diccionario de la Real Academia Española (DRAE) and María Moliner’s electronic version of the Diccionario de uso del español (MM). The derived forms used in this investigation will be affixed words.

The second part of this study is a survey that closely follows the methodology of the experiment by Hay (2001). Using the list of derived forms from the first part of this study, a survey will be created and administered to Spanish-speaking participants. When presented a word pair, the participants will be asked to select the derived form that is more complex, one having a higher base frequency and the other a higher derived-form frequency. The results of the survey will provide insight into the processing of morphologically complex terms.

In addition to the survey, I will also investigate the semantic transparency of the terms. Semantic transparency describes the degree to which the meaning of a morphologically complex word can be derived from the analysis of the component morphemes. For example, abrelatas has a high degree of semantic transparency: it is a tool used to open tin cans. By contrast, comecocos does not have semantic transparency. The meaning ‘obsession or brainteaser’ is not as
predictable from an analysis of the component morphemes. I expect semantic transparency will decrease as the whole-word frequency surpasses the frequency of the base. I will utilize the dictionary of the DRAE and MM to determine if the base is included in the definition of the derived form.

**Part 1 – Word Selection**

The first step was to determine if the study was even practicable. In order to perform the experiment there needed to be enough data to form pairs of derived forms, each pair consisting of a derived form with a higher frequency than its base and a derived form with a lower frequency than its base. Furthermore, enough pairs needed to be compiled so that derived word pairs could be made based on syllable count, stress, phonological patterns, and similar frequencies.

Using the Alameda and Cuetos (1995) corpus, derived words were compiled into lists according to their suffixes or prefixes. The word frequency for each was included as well. In this preliminary phase of the experiment, the following prefixes were used: *a-, an-, ante-, anti-, des-* and *pre-. The following suffixes were selected: *-encia, -ancia, -ente, -tud, -dad, -aje, -ez, -dor* and *-ero*. These prefixes and suffixes were selected from a list of affixes in the textbook *Fuentes* by Tuten, Caycedo Garner and Esterrich (2011) as a starting point to gauge the feasibility of this experiment (pp.123, 129).

The bases of the derived words were then added to the list with their corresponding frequency. The DRAE was used to identify the bases. In this first phase, masculine and feminine forms were considered as well as singular and plural forms. It was later determined that only the masculine singular forms would be included in the final phase due to the lack of sufficient data from the corpus to make meaningful word pairs as well as the need for congruency. In addition,
only the infinitive was used in the final pairings due to the complications of compiling the frequency of a verb with all its inflections and making equivalent word pairs. In this preliminary phase of the study, it was determined that a full experiment would be possible.

**Prefixed Word Selection**

Once it was evident that enough data could be collected, the task was to collect data in an organized and well-founded manner. The selection of prefixed words began with a study of prefixes using Lang (1990). Lang summarizes some of the important differences between prefixes and suffixes (p.168). Some prefixes function independently and are not always considered bound morphemes. For example, some prefixes also function as prepositions and adverbs. Another property of prefixes is that they sometimes display a lower degree of semantic cohesion with their bases; this loose connection may find graphic representation in the form of a connecting hyphen. Lastly, prefixes do not generally alter grammatical class. Lang expands on the characteristics of different types of prefixes and it was this information that shaped the experiment conducted in the present study (Lang, 1990).

Lang (1990) divides prefixes into sense groups based on semantic function. The three groups used in this experiment were prefixes of negation, locative prefixes, and temporal prefixes. The prefixes of negation included *a-/an-, anti-, contra-, des-/dis-/de-, extra*, and *in-/im-/i-* (pp. 170-73). The locative prefixes based on space, position and location consisted of the following: *ante-, entre-, inter-, retro-, sobre-, super-, sub-/so-, and trans-/tras-* (pp. 174-77). The temporal prefixes were *ante-, post-/pos-* and *pre-* (pp. 177-78).

Using the *Find* function of Microsoft Word and changing the options to find prefixes, each prefix was found in Alameda and Cuetos (1995) with a corresponding frequency. Not all the instances of appearance in the dictionary were the prefix, but this search showed that there
were enough occurrences to proceed to the collection of derived words. Table 1 contains a list of each prefix with the number of occurrences found in Alameda and Cuetos (1995) using the *Find* function in Word as explained previously.

**Table 1 Frequency of Selected Prefixes**

<table>
<thead>
<tr>
<th>Negation Prefixes</th>
<th># of occurrences*</th>
<th>Locative Prefixes</th>
<th># of occurrences*</th>
<th>Temporal Prefixes</th>
<th># of occurrences*</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>a-</em></td>
<td>12067**</td>
<td><em>inter-</em></td>
<td>361</td>
<td><em>pre-</em></td>
<td>946</td>
</tr>
<tr>
<td><em>an-</em></td>
<td>1643</td>
<td><em>retro-</em></td>
<td>31</td>
<td><em>ante-</em></td>
<td>61</td>
</tr>
<tr>
<td><em>anti-</em></td>
<td>195</td>
<td><em>sobre-</em></td>
<td>148</td>
<td><em>post-</em></td>
<td>62</td>
</tr>
<tr>
<td><em>contra-</em></td>
<td>181</td>
<td><em>sobre-</em></td>
<td>130</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>des-</em></td>
<td>2702</td>
<td><em>so-</em></td>
<td>1046</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>dis-</em></td>
<td>655</td>
<td><em>sub-</em></td>
<td>293</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>de-</em></td>
<td>4325</td>
<td><em>trans-</em></td>
<td>275</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>extra-</em></td>
<td>118</td>
<td><em>tras-</em></td>
<td>177</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>in-</em></td>
<td>3134</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>im-</em></td>
<td>604</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>i-</em></td>
<td>4491**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*It is important to note that the numbers of occurrences represent not only the prefixes, but also all words that start with those letter combinations.**

**The *Find* feature in Word also counted the occurrences of “a” and “i” used in the column of Alameda and Cuetos to indicate the first letter of each word, so the # of occurrences for “a” or “i” has been divided by 2 because each letter was counted twice.

For each prefix used in the experiment, the general information gathered from *Spanish Word Formation* assisted in the organization of the data collected from Alameda and Cuetos (1995). Each prefix group was entered onto a separate spreadsheet and each prefix in the group was assigned a sheet within the spreadsheet. The derived words of each prefix were organized based on the type of base to which they adhered. For example, according to Lang (1990), the prefix *a-* and the allomorphic variation *an-* tend to adhere to adjective bases and noun bases that are preferably suffixed (p. 170). Thus for the prefix *a-*, the derived words were divided into adjective bases and noun bases, but since the noun bases were suffixed, they were excluded.
Using the DRAE and MM, a list of bases corresponding to the affixed terms was added to each of the Excel spreadsheets. The base frequencies from Alameda and Cuetos (1995) were also included. For the lists contained in the spreadsheets, a logical formula was applied to each derived form and its corresponding base to identify which had the higher frequency (see Fig. 1). Base frequencies that are higher than the whole term frequency are marked BASE (see the center column). If the whole word frequency is higher than the base frequency, DERIVATION is the identifier. For example, *descubrir* has a word frequency of 109 and the verb base *cubrir* has a frequency of 33. This derived word and base would be identified by DERIVATION in the spreadsheet because the derived word has a higher frequency than its base.

The figure below is a sample from the prefix *des*-. Column A shows the general information from Lang (1990) and the number of occurrences in Alameda and Cuetos (1995). Column B contains the base word for the derived word in Column G. Columns C and H show the word frequencies and Column E identifies which has the higher frequency, the base word or the derived word. The blanks in Column E indicate that the base word was not found in Alameda and Cuetos (1995) and therefore no comparison could be made.

*Figure 1. Sample Spreadsheet of the Prefix des*-

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>General Prefix Information from Lang (1990)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td><em>par excellence with a high degree of base amenability</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>affinity to verb bases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>des- productive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2702 matches in Alameda and Cuetos (1995), but not all are the prefix</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>arrollar</td>
<td>1</td>
<td>DERIVATION</td>
<td>desarrollar</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>cubrir</td>
<td>33</td>
<td>DERIVATION</td>
<td>descubrir</td>
<td>109</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>aparecer</td>
<td>90</td>
<td>BASE</td>
<td>desaparecer</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>cansar(se)</td>
<td>2</td>
<td>DERIVATION</td>
<td>descansar</td>
<td>53</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>sacar</td>
<td>0</td>
<td></td>
<td>destacar</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>ayunar</td>
<td>2</td>
<td>DERIVATION</td>
<td>desayunar</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Suffixed Word Selection

The selection of suffixed words also began with a study of suffixes using Lang (1990). Choosing suffixes was more challenging than choosing prefixes. Lang’s discussion of suffixation covers four chapters: “Emotive suffixation”, “Non-emotive suffixation”, “Adjectivization and adverbilization”, and “Verbalisation”. In order to carry out this experiment, it was necessary to limit which suffixes were to be used. The current experiment uses non-emotive suffixes that change the base fundamentally and often change the syntactic category as explained by Lang (p.124). Emotive suffixation, including diminutives, augmentatives and pejoratives, was excluded because while these suffix types alter the base, they tend not to change the grammatical category. The change in grammatical category was extremely useful in selecting the suffixes and organizing the data in order to form word pairs of the same grammatical category.

With the list of suffixes organized into groups based on the change in grammatical category, a search was again performed just like the search that was executed with the prefixes. Using the Find feature of Word for the electronic copy of Alameda and Cuetos (1995) with the advanced option of finding the suffix selected, this search yielded the frequencies for 34 non-emotive suffixes taken from Lang (1990). These frequencies were valuable in narrowing down the 34 non-emotive suffixes to a more feasible number. Six high-frequency suffixes and 6 low-frequency suffixes were selected to be used in the experiment; see Tables 2 and 3, respectively.

Table 2  Frequency of High-frequency Suffixes

<table>
<thead>
<tr>
<th>Suffix</th>
<th>Frequency</th>
<th>Grammatical Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>-al</td>
<td>889</td>
<td>N→N</td>
</tr>
<tr>
<td>-idad</td>
<td>580</td>
<td>ADJ→N</td>
</tr>
<tr>
<td>-dor</td>
<td>495</td>
<td>V→N</td>
</tr>
<tr>
<td>-ero</td>
<td>453</td>
<td>N→N</td>
</tr>
<tr>
<td>-encia</td>
<td>229</td>
<td>ADJ→N</td>
</tr>
<tr>
<td>-ez</td>
<td>121</td>
<td>ADJ→N</td>
</tr>
</tbody>
</table>
Table 3  Frequency of Low-frequency Suffixes

<table>
<thead>
<tr>
<th>Suffix</th>
<th>Frequency</th>
<th>Grammatical Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>-aje</td>
<td>97</td>
<td>N→N</td>
</tr>
<tr>
<td>-ición</td>
<td>80</td>
<td>V→N</td>
</tr>
<tr>
<td>-ancia</td>
<td>65</td>
<td>ADJ→N</td>
</tr>
<tr>
<td>-mento</td>
<td>60</td>
<td>V→N</td>
</tr>
<tr>
<td>-dero</td>
<td>55</td>
<td>V→N</td>
</tr>
<tr>
<td>-dura</td>
<td>53</td>
<td>V→N</td>
</tr>
<tr>
<td>-tud</td>
<td>40</td>
<td>ADJ→N</td>
</tr>
<tr>
<td>-edad</td>
<td>54</td>
<td>ADJ→N</td>
</tr>
<tr>
<td>-tad</td>
<td>33</td>
<td>ADJ→N</td>
</tr>
</tbody>
</table>

Note: Three addition low-frequency suffixes were selected to due to insufficient data to form word pairs.

After the 12 suffixes were selected, derived word lists were created using the same process as with the prefixes, except the suffixes were organized by the change in grammatical category instead of the type of base to which they adhere. Bases for the derived words were found using the DRAE and MM and the frequencies were added using Alameda and Cuetos (1995). Again, the logical formula was applied and each derived word and base were labeled as BASE, to indicate the base frequency was higher, or DERIVATION, to indicate the derived word frequency was higher. The same process was performed for the three addition low-frequency suffixes that were selected in order to yield sufficient data for the low-frequency suffixes.

**Pairing of Derived Words**

Once all the lists of derived words and their corresponding bases were compiled, the pairing of derived forms began. The derived forms were paired by affix and consisted of two derived forms, one labeled as BASE and the other as DERIVATION. A paired example of derived forms is *sociedad* and *suciedad*. *Sociedad* is labeled DERIVATION and *suciedad* is labeled BASE. As far as possible, pairs were created based on similar syllable count, stress, and
phonological patterns. The prefixed pairs were placed in a spreadsheet for prefixes and the suffixed pairs were gathered in a spreadsheet for suffixes. After this step was completed, the pairs were checked using other corpora to verify that the frequency patterns were consistent.

The *Corpus del Español*, which contains 100 million words of text, served as a source for frequency and a check to verify the frequency trends of Alameda and Cuetos (1995). Alameda and Cuetos (1995) is based primarily on texts from Spain whereas the *Corpus del Español* is based two-thirds on text and one-third on spoken Spanish from speakers in 11 different countries (Davies, 2006, p. 2). In order to use a similar time range from the three corpora, the word frequencies from the *Corpus del Español* were limited to the 1900s. In addition to the *Corpus del Español*, the *Google Books Spanish Corpus* of 45 billion words was also used as a source for frequency. Due to the size of the Google corpus, the frequencies were limited to the years 1980 to 2000. These years were selected because they are the most current years in the corpus and better reflect the language used by the survey participants.

Using the additional corpora, the frequencies for each derived word and corresponding base were added into the spreadsheets containing the prefixed and suffixed pairs from the previous step. Figures 2 and 3 are sample spreadsheets. For each derived word pair, the base frequency and the derivation frequency were highlighted to correspond with each frequency source (see Figs. 2 & 3).

In order to verify patterns between the three corpora, each derived word went through the following process for each frequency sources. The derived frequency and base frequency were totaled in a column titled *Token Total*. The derived frequency and the base frequency were each divided by the total number of tokens resulting in a percentage. The resulting percentage served as means to compare the frequency patterns between sources. For the derived forms labeled
DERIVATION to be acceptable, the derived frequency had to account for at least 50% of the token total in all three corpora. For the derived forms labeled BASE to acceptable, the base frequency had to be at least 50% of the token total in all three corpora.

If the frequency percentages for a derived form from the Alameda and Cuetos frequency dictionary, the *Corpus del Español*, and the *Google Books Spanish Corpus* do not agree, the derived form was discarded and another suitable derived-form word pair was used. As an example, *substrato* was discarded because the frequency patterns of the derived form and base from the *Corpus del Español* and the *Google Books Spanish Corpus* did not match the frequency pattern from Alameda and Cuetos (1995) (see Fig. 2). The derived frequency for *substrato* was 78% of the token total based on Alameda and Cuetos (1995), while the percentages for the *Corpus del Español* and the *Google Books Spanish Corpus* were 12.5 and 12.9 respectively.

**Figure 2.** The Prefix *sub-*

<table>
<thead>
<tr>
<th></th>
<th>Derivation - Higher frequency</th>
<th>Token Total</th>
<th>Base - Higher frequency</th>
<th>Token Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rayar</td>
<td>2</td>
<td>0.142357</td>
<td>rayar</td>
<td>12</td>
</tr>
<tr>
<td>Subrayar</td>
<td>12</td>
<td>0.8571</td>
<td>Subrayar</td>
<td>12333535</td>
</tr>
<tr>
<td>Ventir</td>
<td>14</td>
<td>1.3931</td>
<td>Subrayar</td>
<td>12333535</td>
</tr>
<tr>
<td>Subrayar</td>
<td>12</td>
<td>0.8571</td>
<td>Subrayar</td>
<td>12333535</td>
</tr>
<tr>
<td>Locative</td>
<td>2</td>
<td>0.142357</td>
<td>Locative</td>
<td>12333535</td>
</tr>
<tr>
<td>Subrayar</td>
<td>12</td>
<td>0.8571</td>
<td>Subrayar</td>
<td>12333535</td>
</tr>
<tr>
<td>Locative</td>
<td>2</td>
<td>0.142357</td>
<td>Locative</td>
<td>12333535</td>
</tr>
</tbody>
</table>

**Figure 3.** The Suffix *-dor*
After the derived word pairs were verified using the three corpora, the result was 21 prefixed pairs and 27 suffixed pairs for a total of 48 word pairs. In her experiment, Hay (2001) used 34 word pairs, 17 prefixed pairs and 17 suffixed pairs (p. 1046). In addition to the 48 word pairs, 33 pseudo pairs or filler pairs were formed to serve as distractors. The target number of pseudo pairs was 30, the number of filler pairs used by Hay (p. 1047). The pseudo pairs were formed by matching a pseudo-affixed word with an affixed word. An example pair is *visual* and *ritual*, with *visual* being the pseudo-affixed word. The affixed pairs were counterbalanced and randomized with the pseudo pairs. The complete set contained 81 word pairs.

**Part 2 – Survey of Spanish Speakers**

**Pilot Survey**

The pilot survey tested if the experiment by Hay (2001) was possible in Spanish with Spanish speakers. With the completion of the pilot, the word pairs went through a rigorous approval process to narrow down the actual pairs from 48 pairs to 34 pairs. The pilot survey was administered to 11 native Spanish speakers studying Spanish in the Spanish MA program at Brigham Young University. Nine students participated. The instructions for the pilot survey followed closely those of Hay (2001), although they were in Spanish with examples of Spanish words (p. 1048). In addition to the explanation of complex words, an example was included after the explanation (see Appendix C).

The results of the pilot survey were varied. The results for two of the nine participants were discarded due to the high percentage of the pseudo-affixed word being selected. In the two discarded surveys, participants chose the pseudo-affixed word for 64% and 70% of the pseudo pairs. For the remaining seven surveys, the percentages were 12, 18, 21, 42, 48, 48 and 51. Three of the survey responses were considered valid and the remaining four responses were
questionable due to the percentage of the pseudo-affixed word being selected in the pairs that served as a distraction. The analysis of the data showed that the derived word with the higher-frequency base was selected 56% of the time. The derived word with the lower-frequency base was selected 44% of the time. This was promising as it showed a slight indication that relative frequency does indeed play a role in decomposition.

The results of the survey did not prove useful in eliminating the extra 14 pairs. In order to eliminate the extra derived word pairs, the following process was employed. First, the word pairs were separated into affixed pairs and pseudo pairs. The word pairs were separated back into two groups by affix: 21 prefixed pairs and 27 suffixed pairs. Pairs were eliminated based on the following criteria: word structure, the relationship of derived frequency and base frequency, and the relationship of the two derived frequencies of each pair.

The word structure of each pair was closely examined based on syllable count, stress and phonological patterns. The pair ingeniero and invernadero was eliminated on account of syllable count. Each derived word was also analyzed by its component morphemes and its categorization into the two groups: prefixed pairs and suffixed pairs. Three prefixed pairs were eliminated. While both words of the pair contained a prefix, the base of the affixed word also contained the prefix thus creating confusion on how the word was decomposed. For example, predominante was eliminated. While predominante can be separated into pre- and dominante, the base of predominante is predominar. A fourth prefixed pair was also eliminated because one of the words of the pair also exemplified this issue. The suffixed pair obediencia and evidencia was also eliminated due to the diphthong in the penultimate syllable of obediencia and the lack of the same diphthong in evidencia.
The relationship of derived frequency and base frequency of each derived word was verified and no pairs were eliminated. Column E had to be .5 or greater indicating the derived word frequency was higher than the base frequency (see Fig. 4). Column L had to be .5 or greater for the derived words whose base frequency was higher than the derivation frequency.

Figure 4. View of Excel Spreadsheet with Word A, the Derived Form with a Higher Derived Frequency, Compared to Word B, the Derived Form with a Higher Base Frequency

<table>
<thead>
<tr>
<th>Word A</th>
<th>freq</th>
<th>base freq</th>
<th>Derived</th>
<th>Base</th>
<th>Token sum</th>
<th>Word B</th>
<th>freq</th>
<th>base freq</th>
<th>Derived</th>
<th>Base</th>
<th>Token sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>desayuno</td>
<td>62</td>
<td>2</td>
<td>0.969</td>
<td>0.031</td>
<td>54</td>
<td>descanso</td>
<td>16</td>
<td>44</td>
<td>0.267</td>
<td>0.733</td>
<td>60</td>
</tr>
<tr>
<td>increíble</td>
<td>50</td>
<td>4</td>
<td>0.926</td>
<td>0.074</td>
<td>54</td>
<td>invisible</td>
<td>54</td>
<td>57</td>
<td>0.486</td>
<td>0.514</td>
<td>111</td>
</tr>
<tr>
<td>infinito</td>
<td>73</td>
<td>15</td>
<td>0.820</td>
<td>0.180</td>
<td>89</td>
<td>indirecto</td>
<td>7</td>
<td>53</td>
<td>0.117</td>
<td>0.883</td>
<td>60</td>
</tr>
<tr>
<td>inagotable</td>
<td>13</td>
<td>1</td>
<td>0.929</td>
<td>0.071</td>
<td>14</td>
<td>inacceptable</td>
<td>6</td>
<td>18</td>
<td>0.250</td>
<td>0.750</td>
<td>24</td>
</tr>
<tr>
<td>inaudible</td>
<td>4</td>
<td>4</td>
<td>0.500</td>
<td>0.500</td>
<td>8</td>
<td>intangible</td>
<td>4</td>
<td>11</td>
<td>0.267</td>
<td>0.733</td>
<td>15</td>
</tr>
<tr>
<td>inseparable</td>
<td>23</td>
<td>1</td>
<td>0.958</td>
<td>0.042</td>
<td>24</td>
<td>inseparable</td>
<td>11</td>
<td>35</td>
<td>0.239</td>
<td>0.761</td>
<td>40</td>
</tr>
<tr>
<td>interlocutor</td>
<td>33</td>
<td>30</td>
<td>0.524</td>
<td>0.476</td>
<td>63</td>
<td>intercelular</td>
<td>2</td>
<td>80</td>
<td>0.022</td>
<td>0.978</td>
<td>91</td>
</tr>
<tr>
<td>irresoluto</td>
<td>6</td>
<td>1</td>
<td>0.857</td>
<td>0.143</td>
<td>7</td>
<td>irresponsable</td>
<td>9</td>
<td>74</td>
<td>0.108</td>
<td>0.892</td>
<td>83</td>
</tr>
<tr>
<td>intender</td>
<td>2</td>
<td>1</td>
<td>0.667</td>
<td>0.333</td>
<td>3</td>
<td>irpetido</td>
<td>1</td>
<td>29</td>
<td>0.033</td>
<td>0.967</td>
<td>30</td>
</tr>
<tr>
<td>pretender</td>
<td>25</td>
<td>8</td>
<td>0.738</td>
<td>0.242</td>
<td>33</td>
<td>preceder</td>
<td>2</td>
<td>24</td>
<td>0.077</td>
<td>0.923</td>
<td>25</td>
</tr>
<tr>
<td>irremediable</td>
<td>4</td>
<td>3</td>
<td>0.571</td>
<td>0.429</td>
<td>7</td>
<td>irremediable</td>
<td>8</td>
<td>11</td>
<td>0.421</td>
<td>0.579</td>
<td>19</td>
</tr>
<tr>
<td>subrayar</td>
<td>12</td>
<td>1</td>
<td>0.925</td>
<td>0.077</td>
<td>13</td>
<td>subvenir</td>
<td>1</td>
<td>123</td>
<td>0.008</td>
<td>0.992</td>
<td>124</td>
</tr>
<tr>
<td>sobresaliente</td>
<td>5</td>
<td>3</td>
<td>0.625</td>
<td>0.375</td>
<td>8</td>
<td>sobreviendo</td>
<td>2</td>
<td>12</td>
<td>0.143</td>
<td>0.857</td>
<td>14</td>
</tr>
<tr>
<td>transportar</td>
<td>11</td>
<td>3</td>
<td>0.796</td>
<td>0.214</td>
<td>14</td>
<td>transformar</td>
<td>31</td>
<td>57</td>
<td>0.352</td>
<td>0.648</td>
<td>88</td>
</tr>
</tbody>
</table>

Finally, the relationship of the derived frequencies for each word pair was vital in eliminating pairs. Only those pairs whose frequencies were closest to a 1:1 ratio were included. The ratio was calculated by dividing the derived frequency of the derived word whose base frequency was lower by the derived frequency of the derived word whose base frequency was higher. This action kept the derived pairs with derived frequencies closer to a 1:1 ratio, thus controlling for the influence of absolute frequency and allowing the influence of relative frequency to be seen. Pairs were eliminated based on the distance from 0 after calculating the log of the ratio, which sets the 1:1 ratio at 0. Before the elimination process based on the criteria of the derived frequency relationship, there were 16 prefixed pairs and 26 suffixed pairs.

The derived prefixed pair with a log of 1.76 was eliminated because it was the pair furthest from 0 (see Fig. 4). The final prefixed pairs were within the range -.45 and 1.22 for the
log of the derived frequency ratio (see Fig. 5). More pairs could have been eliminated, but in order to keep sufficient pairs for the survey, the 15 pairs with a log closest to 0 were kept.

*Figure 5. Graph Depicting the Log of the Ratio of Derived Frequencies of Prefixed Pairs*

The graph in Figure 6 represents the pairs kept for the final survey.

*Figure 6. Graph of Log of Ratio of Derived Frequencies of 15 Prefixed Pairs for Survey*

For the suffixed pairs, the 15 pairs closest to the 1:1 ratio of derived frequencies were kept. As with the prefixed pairs, the log of the ratio was calculated and the pairs farthest from 0
were eliminated. Pairs with a log greater than .45 or less than -.45 were eliminated, resulting in 15 pairs remaining for the survey. This range was tighter than that of the prefixed pairs due to the higher number of pairs that needed to be eliminated. Eleven suffixed pairs were eliminated versus the one prefixed pair that was eliminated; see Figs. 7 and 8.

Figure 7. Graph Depicting the Log of the Ratio of Derived Frequencies of Suffixed Pairs

![Graph Depicting the Log of the Ratio of Derived Frequencies of Suffixed Pairs](image1)

Figure 8. Graph of Log of Ratio of Derived Frequencies of 15 Suffixed Pairs for Survey

![Graph of Log of Ratio of Derived Frequencies of 15 Suffixed Pairs for Survey](image2)
Tables 4 and 5 represent the 15 prefixed pairs and the 15 suffixed pairs remaining after the elimination process based on the criteria of the relationship of derived frequencies.

**Table 4  Finalized Prefixed Pairs in the Survey**

<table>
<thead>
<tr>
<th>Word A</th>
<th>freq.*</th>
<th>base freq.</th>
<th>Derived - % of Token Total</th>
<th>Token Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>sobrecogedor</td>
<td>7</td>
<td>0</td>
<td>1.000</td>
<td>7</td>
</tr>
<tr>
<td>desayuno</td>
<td>62</td>
<td>2</td>
<td>0.569</td>
<td>64</td>
</tr>
<tr>
<td>increíble</td>
<td>50</td>
<td>4</td>
<td>0.526</td>
<td>54</td>
</tr>
<tr>
<td>infinito</td>
<td>73</td>
<td>16</td>
<td>0.820</td>
<td>89</td>
</tr>
<tr>
<td>insaciable</td>
<td>13</td>
<td>1</td>
<td>0.929</td>
<td>14</td>
</tr>
<tr>
<td>inaudible</td>
<td>4</td>
<td>4</td>
<td>0.500</td>
<td>8</td>
</tr>
<tr>
<td>inmutable</td>
<td>23</td>
<td>1</td>
<td>0.558</td>
<td>24</td>
</tr>
<tr>
<td>interlocutor</td>
<td>33</td>
<td>30</td>
<td>0.524</td>
<td>63</td>
</tr>
<tr>
<td>irrefrangible</td>
<td>6</td>
<td>1</td>
<td>0.857</td>
<td>7</td>
</tr>
<tr>
<td>irresoluto</td>
<td>2</td>
<td>1</td>
<td>0.667</td>
<td>3</td>
</tr>
<tr>
<td>pretender</td>
<td>25</td>
<td>8</td>
<td>0.758</td>
<td>33</td>
</tr>
<tr>
<td>irrefreniente</td>
<td>4</td>
<td>3</td>
<td>0.571</td>
<td>7</td>
</tr>
<tr>
<td>subrayar</td>
<td>12</td>
<td>1</td>
<td>0.523</td>
<td>13</td>
</tr>
<tr>
<td>sobresaliente</td>
<td>5</td>
<td>3</td>
<td>0.625</td>
<td>8</td>
</tr>
<tr>
<td>trasportar</td>
<td>11</td>
<td>3</td>
<td>0.788</td>
<td>14</td>
</tr>
</tbody>
</table>

**Total 15 Pairs**

Word A - derived word has a higher frequency
Word B - base has a higher frequency
*Frequencies based on Alameda and Cuetos (1995)

**Table 5  Finalized Suffixed Pairs in the Survey**

<table>
<thead>
<tr>
<th>Word A</th>
<th>freq.*</th>
<th>base freq.</th>
<th>Derived - % of Token Total</th>
<th>Token Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>amplificador</td>
<td>7</td>
<td>1</td>
<td>0.875</td>
<td>6</td>
</tr>
<tr>
<td>aventurero</td>
<td>23</td>
<td>13</td>
<td>0.639</td>
<td>36</td>
</tr>
<tr>
<td>ambigüedad</td>
<td>46</td>
<td>31</td>
<td>0.597</td>
<td>77</td>
</tr>
<tr>
<td>habilidad</td>
<td>52</td>
<td>15</td>
<td>0.776</td>
<td>67</td>
</tr>
<tr>
<td>asido</td>
<td>5</td>
<td>2</td>
<td>0.714</td>
<td>7</td>
</tr>
<tr>
<td>fundador</td>
<td>17</td>
<td>13</td>
<td>0.567</td>
<td>30</td>
</tr>
<tr>
<td>actividad</td>
<td>240</td>
<td>56</td>
<td>0.811</td>
<td>296</td>
</tr>
<tr>
<td>relevancia</td>
<td>8</td>
<td>1</td>
<td>0.889</td>
<td>9</td>
</tr>
<tr>
<td>edición</td>
<td>53</td>
<td>4</td>
<td>0.930</td>
<td>57</td>
</tr>
<tr>
<td>labrador</td>
<td>9</td>
<td>1</td>
<td>0.900</td>
<td>10</td>
</tr>
<tr>
<td>creatividad</td>
<td>22</td>
<td>11</td>
<td>0.667</td>
<td>33</td>
</tr>
<tr>
<td>nutrición</td>
<td>9</td>
<td>1</td>
<td>0.900</td>
<td>10</td>
</tr>
<tr>
<td>magnitud</td>
<td>43</td>
<td>8</td>
<td>0.843</td>
<td>51</td>
</tr>
<tr>
<td>orador</td>
<td>6</td>
<td>2</td>
<td>0.750</td>
<td>8</td>
</tr>
<tr>
<td>profesional</td>
<td>103</td>
<td>69</td>
<td>0.599</td>
<td>172</td>
</tr>
</tbody>
</table>

**Total 15 Pairs**

Word A - derived word has a higher frequency
Word B - base has a higher frequency
*Frequencies based on Alameda and Cuetos (1995)
After narrowing down the affixed pairs from the 48 pairs in the pilot survey to 30 total pairs for the actual survey, the 33 filler pairs in the pilot were also examined. Eighteen of the 33 pairs were kept without changes. Seven pairs had one of the words in the pair changed because the filler affixed word was not a clear representation of the desired concept. The pseudo affixed words needed to be words that did not contain an apparent base. Eight pairs were excluded because they were selected more than three out of seven times in the pilot survey, thus indicating that participants viewed the pseudo word as being decomposable. Five new pairs were added to the stimuli to achieve 15 filler prefixed pairs and 15 filler suffixed pairs for a total of 30 filler pairs in the final survey. The filler pairs in Table 6 served as distractors in the final survey.

Table 6 Finalized Filler Pairs in the Survey

<table>
<thead>
<tr>
<th>15 Filler Prefixed Pairs</th>
<th>15 Filler Suffixed pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>antecesor</td>
<td>ambición</td>
</tr>
<tr>
<td>anticipo</td>
<td>carpintero</td>
</tr>
<tr>
<td>contradictor</td>
<td>estadístico</td>
</tr>
<tr>
<td>delatar</td>
<td>inteligente</td>
</tr>
<tr>
<td>desenfilar</td>
<td>majestad</td>
</tr>
<tr>
<td>despejar</td>
<td>mamífero</td>
</tr>
<tr>
<td>disparete</td>
<td>oriente</td>
</tr>
<tr>
<td>disyuntivo</td>
<td>paciente</td>
</tr>
<tr>
<td>imperio</td>
<td>posición</td>
</tr>
<tr>
<td>preservar</td>
<td>secuencia</td>
</tr>
<tr>
<td>prevalecer</td>
<td>serpiente</td>
</tr>
<tr>
<td>prevención</td>
<td>superstición</td>
</tr>
<tr>
<td>retrospectiva</td>
<td>virtud</td>
</tr>
<tr>
<td>sublevar</td>
<td>visual</td>
</tr>
<tr>
<td>superficie</td>
<td>voluntad</td>
</tr>
<tr>
<td>superponer</td>
<td>superego</td>
</tr>
</tbody>
</table>

In addition to the analysis of the affixed and pseudo pairs, clarity was added to the final survey instructions in order to increase the success of the survey. Instead of just a single example at the conclusion of the instructions, three examples were provided. Some examples in the explanation were also changed to better reflect the concept of complex words as explained in the instructions.
Final Survey

With all the details worked out from the pilot survey, the next part of the experiment could begin. This second part of the experiment is the survey on the decompositionality of morphologically complex terms. The survey explores the effects of frequency and how speakers of Spanish view affixed terms in terms of ease of decomposition. The survey consists of the 60 word pairs from the first part of the experiment, 30 affixed pairs and 30 filler pairs. Each affixed pair consists of a derived word whose base frequency is higher than the derived word frequency and a derived word with a higher whole-word frequency than the base frequency. It is anticipated that the derived word with the higher base frequency will be selected as the morphologically complex term that is easier to decompose. It is also hypothesized that when the base frequency is higher than the derived word frequency, the more likely decomposition occurs.

The survey was created in Qualtrics, a web-based tool for building surveys, and a link was emailed to participants. The preliminary questions of the survey contain demographic questions thereby allowing the responses to be categorized into native speaker responses and Spanish learner responses. The responses to the demographic questions will provide information about the time spent speaking and learning Spanish and the regional variety with which each participant identifies.

Participants were provided with instructions containing a brief explanation of complex terms and a few examples. The instructions closely followed those of Hay (2001), although they are in Spanish with examples of Spanish words. Participants were asked to read each derived word pair silently and select the word they found more complex. After acknowledging that they understood the instructions, participants were directed to the finalized survey pairs in which they selected which word of the 60 pairs was easier to break into more meaningful parts. There was
no time limit to the survey, allowing participants to move at their own pace. The survey was intended to take 10 to 15 minutes; see Fig. 9 for the survey instructions.

Figure 9. Survey Instructions

Cuestionario sobre palabras complejas

Lo siguiente es un experimento con palabras complejas.

Una palabra compleja es una palabra que se puede dividir en unidades más pequeñas que tienen significado.

En español, por ejemplo, la palabra crueldad se divide en dos partes: cruel y -dad. -dad es una unidad que aparece al final de muchas palabras españolas. En crueldad, -dad ha sido añadido a cruel para formar la nueva palabra crueldad, que es más compleja. Una palabra que está compuesta por unidades más pequeñas así es una palabra compleja.

Contraluz es otro ejemplo de una palabra compleja en español. Se puede separar en dos partes, contra- y luz.

Las palabras que no son complejas son simples. Algunos ejemplos de palabras simples en español son mesa, amarillo, o arte. Es imposible dividir la palabra mesa en unidades más pequeñas. Masa no es compleja.

En el siguiente experimento, se presentarán pares de palabras complejas. Los participantes deben decidir cuál de las dos palabras del par es la MÁS compleja.


Otro ejemplo de una palabra compleja es negritud. Aunque una persona nunca haya escuchado esta palabra antes, podría entender el significado, porque se puede dividir en negro y -itud. Gratitud también es compleja – se puede dividir en grato y -itud. Pero gratitud no parece tan compleja como negritud. No necesitamos dividir gratitud en partes para entender el significado, y, de hecho, no parece natural hacerlo.

Para cada uno de los siguientes pares de palabras, tenga la bondad de leer ambas palabras en voz baja, y luego seleccione la palabra más compleja, la que es más fácil de dividir en partes que tienen significado. Es muy importante seleccionar una palabra de cada par, aunque Ud. no esté seguro. No deje ninguna respuesta en blanco. Si es necesario, siga su intuición y haga la mejor selección que pueda.

Ejemplos

- negritud
- gratitud
- desenlar
- desmontar
- sucesidad
- sociedad
Participants

The survey was distributed to native Spanish speakers enrolled in the English Language Center (ELC) at Brigham Young University (BYU). In addition to the students at the ELC, the survey was also distributed to Spanish students at BYU enrolled in a 300- or 400-level Spanish class. Participation was voluntary and anonymous. The two groups of participants were chosen based on their Spanish abilities and as a means to compare the morphological processing of native speakers of Spanish to the morphological processing of L2 Spanish speakers.

Nineteen native Spanish speakers participated in the survey, 17 of which were enrolled in the English language program at BYU and two of which were enrolled in a 300- or 400-level class but identified their native language as Spanish. Three of the responses were excluded due to incompletion of the online survey. Of the remaining 16 participants, six identified their regional variety as Mexican Spanish. Three identified Peruvian Spanish as their regional variety. Other regional varieties identified included Uruguayan, Colombian, Venezuelan, Argentinian, Honduran, and Puerto Rican, with each variety being represented by a single participant. Fourteen of the 16 native speakers had received formal education in their regional variety. Two participants received less than four years of education while the rest of the participants had received 11-24 years of formal education in their regional variety. The two native speakers enrolled in the Spanish classes had not received any formal education in their regional variety. All native Spanish-speaking participants have spoken their regional variety for 19-34 years.

Twenty-six L2 Spanish speakers participated in the survey, all enrolled in a 300- or 400-level Spanish class at BYU. Twenty-two of the 26 participants had spent time in a Spanish-speaking country. The remaining four participants indicated that they had at least two years of contact with the Spanish language. Twenty-one of the 22 participants who had spent time in a
Spanish-speaking country had spent more than 1.5 years in that country with the majority spending two years in a Spanish-speaking country. Thirteen participants had spent time in Mexico, eight of which had spent two years in the country. Three participants had spent time in Chile, two who were in the country for 1.5 years and one who was there for two years. The following is a breakdown of the number of participants and the country in which they spent time: two in Guatemala, two in Honduras, two in Spain, two in Peru, one in the Dominican Republic, one in El Salvador, one in Costa Rica, and one in Argentina. Participants had spent time in 10 different countries.

L2 Spanish speakers were also asked to describe the amount of contact they had with the various regional varieties of Spanish. The list of regional varieties coincided with Spanish-speaking countries. Participants listed multiple regional varieties and that contact lasted from anywhere from one week to four years. The only regional variety not selected by any participant was Panamanian Spanish. The majority had contact with Mexican Spanish, which is not surprising given that 13 of the participants had spent time in Mexico.
Chapter 4 – Results

Results from the survey

Eleven of the 16 native Spanish-speaker responses were included in the analysis. The two native speakers enrolled in a Spanish class were separated from the native speaker group because they had not received any formal education in their native language and they were enrolled in a Spanish class unlike the rest of the native speakers. These two participants’ responses were analyzed separately. Three additional native speaker responses were excluded because they selected the filler word more than 20 times out of 30 pairs.

Hay (2001) excluded “subjects who did not provide the same answer (in either direction) for at least 20 of the 30 fillers” (p. 1049). While she reversed answers that were believed to be caused by confusion of a terminological nature, no answers were reversed for this experiment. Subjects that selected the filler word more than 20 times were excluded. Hay’s results (2001) did not show significant difference in the decomposition of prefixed words versus suffixed words and therefore in this experiment the affixed words were not separated for the analysis.

For the 11 native Spanish-speaker responses analyzed, no significant difference was found in the selection of derived forms. Derived forms with a more frequent base were not selected as more complex than the derived forms with a less frequent base. A Wilcoxon ranked sum test was performed (Wilcoxon, by subjects: \( W = 116.5, p = .505 \), by item: \( W = 901, p = .834 \)). In addition, a chi-squared test was run for the responses of the nine participants who chose the filler word less than 15 times of 30 pairs to see if there was a difference. The results again showed no significant difference between selection of the derived form with a higher base frequency and the derived form with a lower base frequency (\( \chi^2(1) = .015, p = .9031 \)). The derived form with a higher base frequency was chosen 134 times while the derived form with a
lower base frequency was chosen 136 times. While the results of the pilot survey showed an indication that relative frequency might play a role in decomposition of derived forms, the results were not statistically significant ($x^2(1) = 2.756, p = .0969$). In contrast, a chi-squared test performed on the responses of the two native speakers enrolled in a 300- or 400- level Spanish class showed that the frequency of the base form plays a role in decomposition ($x^2(1) = 11.267, p = .0008$). The results were statistically significant for the two native Spanish-speakers in a Spanish class.

The results for the L2 Spanish speakers were very different from the results of the native speakers. No participant chose the filler word more than 15 times. In fact, the filler was chosen an average of 21% for all 26 participants. Because the filler was not chosen often by the L2 Spanish speakers, no responses were excluded unlike the native Spanish-speaker sample. A Wilcoxon ranked sum test was also performed for the L2 Spanish-speaker sample (Wilcoxon, by subjects: $W = 379, p < .0001$, by item: $W = 555, p < .001$). These results were statistically significant. The L2 Spanish speakers chose the derived form with a higher base frequency 63% of the time. They chose the derived form with the lower frequency base 37% of the time. These results resemble the results from the experiment by Hay (2001) and provide evidence that the relative frequency of the derived form and base play a role in decomposition for L2 Spanish speakers.

While the results of this survey do not provide evidence that the base frequency of a derived form influences decomposition for native Spanish speakers, the results do provide significant evidence that the base frequency of a derived form does influence decomposition for L2 Spanish speakers (See Table 7). Although the responses from the native Spanish speakers enrolled in a Spanish class do provide convincing evidence that native speakers without any
formal education in their native language tend to view derived forms with a higher frequency base as more complex, more responses are needed to substantiate any conclusions.

Table 7 Survey Results of Wilcoxon Test

<table>
<thead>
<tr>
<th></th>
<th>By Subject</th>
<th></th>
<th>By Item</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Native Speakers</td>
<td>W = 116.5</td>
<td>𝑝 = .505</td>
<td>W = 901</td>
<td>𝑝 = .834</td>
</tr>
<tr>
<td>L2 Speakers</td>
<td>W = 379</td>
<td>𝑝 &lt; .001</td>
<td>W = 555</td>
<td>𝑝 &lt; .001</td>
</tr>
</tbody>
</table>

There was not enough data to compare results by regional variety. Mexican Spanish was the variety selected most by participants in the sample of native Speakers and the results from the native Mexican Spanish speakers resembled the results of the complete native speaker analysis. The results from the L2 Spanish speakers who spent time in Mexico were representative of the responses from the L2 Spanish-speaker analysis. No differences were noted by regional variety though a larger sample size is needed to be conclusive.

Discussion

The results of the native Spanish speakers do not provide evidence that relative frequency, between the derived form and base form, affects the decomposition of affixed words in Spanish. Derived words were paired based on the surface frequencies having a ratio closest to 1:1. This controlled for absolute frequency influencing the responses. The results showed that native speakers were just as likely to choose the derived word with a higher base frequency as they were to choose the derived word with a lower base frequency, suggesting that the derived words were similarly complex. The role of frequency is unclear in the decomposition of derived words such as those used in this experiment. Since the surface frequencies were similar and the derived form with the higher base frequency was chosen nearly as often as the derived form with the lower base frequency, it may be surface frequency that plays a role in decomposition for native Spanish speakers as research suggests.
On the other hand, the results of L2 Spanish speakers show an influence of relative frequency in decomposition, a characteristic not shared with native Spanish speakers. The results are quite distinct from the native Spanish-speaker sample. This study provides evidence that native Spanish speakers and L2 Spanish speakers view morphologically complex terms differently. This raises the question as to why the results are so different.

The reason behind the difference may be related second-language learning. While the native Spanish speakers have a depth of experience and a larger lexicon, the L2 learners lack the same experience with the language and have a smaller L2 lexicon. This experience with the language translates to frequency.

The L2 learners also have the influence of L1 as they process the L2 language. Because L2 learners are not as well versed in the language, they may employ other strategies to increase comprehension. L2 learners may notice similarities with their L1 language, such as cognates. Many of the words in this study are English cognates. The influence of English may be why the results of the L2 learners resembled the results of the study by Hay (2001).

For the L2 learners, the higher frequency base was more visible, thus indicating that they follow a decomposed route when the base frequency is higher than the surface frequency of a derived form. The L2 learners are more likely to follow a dual-route for processing morphologically complex words. When the higher frequency base is more visible, decomposition may occur, whereas a direct-route process is preferred when the whole-word frequency is higher than the base. Native speakers showed no difference in processing strategies, suggesting a single-route model.

The difference in selection of the filler between the native speakers and the L2 learners prompted a closer examination of the filler pairs. For example, in the filler pair carpintero and
carnicero, the native Spanish-speaking participants selected carpintero as easier to decompose. Eight of the 11 native participants chose the pseudo affixed word. Only four of the 26 L2 Spanish speakers chose carpintero. In contrast, 15 of the 26 L2 Spanish speakers chose mamífero as being more complex than mortífero. Only four of the native speakers chose mamífero. It is my belief that the native Spanish speakers saw muerte. Native Spanish speakers were able to recognize related words because of their larger lexicon while it is more probable that L2 Spanish speakers analyzed the derived words for meaning and only recognized bases that helped them understand the derived word.

**Semantic Transparency Analysis**

Sixteen of the 30 word pairs contained the base in the definition of both derived words. The DRAE and MM were used to determine if the base was included in the definition of the derived form. The native Spanish speakers selected the derived word with the lower base frequency nine out of the 16 word pairs. For four of the pairs, the derived word with the lower base frequency was selected just as much as the derived word with the higher base frequency. The L2 learners chose the derived word with the higher frequency base for 13 of the 16 pairs.

Nine of the 30 word pairs contained the base in the definition of one of the derived words. While this difference in semantic transparency might lead to a difference in decomposition, the native speakers continued to select the derived word with a lower frequency base even if that word did not contain the base in the definition. The L2 learners continued to select the derived form with a higher frequency base even when the base was not included in the definition.

Five of the word pairs did not contain the base in the definition of either word. The L2 learner continued to select the derived word with the higher frequency base. The native speakers
chose the derived form with the higher frequency base one more time than the derived form with the lower frequency base.

Fourteen of the 30 word pairs contained at least one derived word without the base in the definition. It is interesting to note that of the 30 derived words with a lower frequency base used in the survey, 12 words did not have the base included in the definition. Only seven of the 30 derived words with a higher frequency base did not have the base included in the definition. This may be coincidence, but it is more likely that the higher the base frequency of a derived form, the more likely semantic transparency increases, especially for L2 learners. While only 60 words were used for this analysis, it appears that semantic transparency does decrease as the whole-word frequency surpasses the frequency of the base. For a more detailed look at semantic transparency and decomposition, see Hay (2001).

**Summary**

Relative frequency does in fact influence decomposition of derived words in Spanish for L2 Spanish speakers. L2 Spanish speakers rated derived words that had a higher frequency base as being more complex, indicating that these derived words were easier to decompose than their counterparts with a lower frequency base. This suggests a dual-route model is more likely for L2 Spanish speakers. Derived words are decomposed when the base frequency of a derived word exceeds the surface frequency, while derived words are processed whole when the surface frequency is higher than the base frequency.

By contrast, relative frequency appears not to influence decomposition for native Spanish speakers. Native Spanish speakers did not indicate a significant difference in the complexity of the derived word pairs. The results suggest that native Spanish speakers employ a single-route model, processing derived words with a higher frequency base in the same way that they process
derived words with a lower frequency base. Both the results of the pilot survey and the results of the survey lack evidence that relative frequency plays a significant role in the decomposition of derived words for native Spanish speakers. Not only do L2 Spanish speakers use different processes based on frequency, but native and L2 learners process the language differently.
Chapter 5 – Conclusion

Research Questions

The first research question was, what role does frequency play in the decomposition of Spanish affixed words? My results indicate that frequency plays a role in the decomposition of Spanish affixed words from the perspective of an L2 Spanish speaker. L2 Spanish speakers tend to decompose a derived form that has a higher frequency base. In the survey portion of this study, L2 Spanish speakers selected the derived form with a higher base frequency as easier to decompose.

For native Spanish speakers, relative frequency did not play a major role in the decomposition of derived words. Surface frequency may play a greater role though further research is needed to support such a claim. Since only the pairs with surface frequencies closest to a 1:1 ratio, the exact role of surface frequency in this experiment in unknown.

As for the second question (i.e., do the same trends found in Hay’s research (2001) hold true for Spanish?), the same trends do not hold true for native Spanish speakers, but they do hold true for L2 Spanish speakers. The results of the L2 Spanish-speaker sample resembled the results of Hay (2001) whereas the results of the native Spanish speakers do not follow the trends of the research by Hay (2001). Native Spanish speakers did not rate derived forms with a higher base frequency as more complex and easier to decompose.

The final question was, does semantic transparency affect decomposition? Of the 60 derived words used in this experiment, 12 words that had a lower frequency base lacked the base in the definition whereas only seven derived words with a higher frequency base lacked the base in the definition. While the appearance of the base in the definition did not seem to affect native Spanish-speaker responses, it may have affected the responses of the L2 learners because relative
frequency played a role in decomposition. A more thorough investigation is needed to explore the effect of semantic transparency on decomposition. An analysis like that of Hay (2001) would prove useful for Spanish.

**Research Implications**

Further research is needed to investigate the role of relative frequency in the decomposition of morphologically complex Spanish words from the perspective of native Spanish speakers. It was quite surprising that the results from the native speakers did not mirror the results of Hay (2001). This may be due to the fact that English is a more analytical language than Spanish. A further examination of the effects of frequency on decomposition of Spanish derived words would be helpful in understanding the decomposition process for native speakers.

The results of L2 Spanish speakers provide insight into second language acquisition and have implications for teaching. Based on the results of this study, learning high-frequency words is very helpful in comprehension of derived words as participants in this study viewed the base in the process of decomposition. Instead of memorizing lists of word, students should be instructed on the components of words in order to discover connections and generalizations.

While this study focused on the relative frequency of a base and a derived word, the base occurs in other words as well. This experiment only considered the frequency of the base and the surface frequency of the derived word. De Jong, Schreuder, and Baayen (2000) discuss other frequency counts, such as Base Frequency or the summed frequency, and the Cumulative Root Frequency that is the sum of Base Frequency and the Family Frequency, which is the total frequency of related morphological family members (p. 329). In order to better understand the role of relative frequency, these other frequency counts should be considered as well.
Appendix A – Consent forms

The following is the consent form used in the pilot survey.

Implied Consent

My name is Guinevere Deaver and I am a graduate student in Hispanic Linguistics at Brigham Young University conducting this research under the supervision of Jeffrey Turley, Associate Professor of Spanish and Portuguese at BYU. You are invited to participate in this research study of the effects of word frequency on the processing of the complex Spanish words.

Your participation in this study will require the completion of the attached survey. This should take approximately 10 minutes of your time. Your participation will be anonymous and you will not be contacted again in the future. You will not be paid for being in this study. This survey involves minimal risk to you. The benefits, however, may impact society by helping increase knowledge about how Spanish speakers process complex words.

You do not have to participate in this study if you do not want to participate. You do not have to answer any question that you do not want to answer for any reason. We will be happy to answer any questions you have about this study. If you have further questions about this project or if you have a research-related problem, you may contact me, Guinevere Deaver, at handdeaver@byu.edu, or my advisor, Dr. Jeffrey Turley at jeffrey_turley@byu.edu.

If you have any questions about your rights as a research participant, you may contact the IRB Administrator at A-285 ASB, Brigham Young University, Provo, UT 84602; irb@byu.edu; (801) 422-1461. The IRB is a group of people who review research studies to protect the rights and welfare of research participants.

The completion of this survey implies your consent to participate. If you choose to participate, please complete the attached survey and email it back to jasmine_talbot@byu.edu by October 18, 2013. Thank you!

Consentimiento Implicito

Este proyecto de investigación lo dirige Guinevere Deaver, estudiante de posgrado de lingüística hispánica del Departamento de Español y Portugués de la Universidad Brigham Young. El Dr. Jeffrey Turley, Profesor del Departamento de Español y Portugués de BYU, supervisa la investigación. Lo invitamos a Ud. a participar en esta investigación sobre los efectos de las frecuencias de palabra en el procesamiento de las palabras complejas del español.

Su participación en el estudio consiste en completar el cuestionario adjunto, el cual requiere aproximadamente 10 minutos para completarse. Su participación será anónima y no se le volverá a contactar en el futuro. No recibirá retribución alguna por participar en el estudio. Los riesgos de participar en el estudio son mínimos. Sin embargo, los beneficios podrán afectar a la sociedad al aumentar el conocimiento del procesamiento mental de las palabras complejas del español.

Su participación en este estudio es voluntaria. Ud. no está obligado a participar en este estudio si no quiere. No tiene que contestar ninguna pregunta si por alguna razón no quiere hacerlo. Estamos dispuestos a contestar cualquier pregunta en cuanto al estudio. Si tiene alguna pregunta en cuanto al estudio o si tiene algún problema con la investigación, podría contactarme a mí, Guinevere Deaver, handdeaver@byu.edu, o a mi supervisor, Dr. Jeffrey Turley, jeffrey_turley@byu.edu.

Si tiene alguna pregunta en cuanto a sus derechos como participante de investigación, podría contactar al Administrador del IRB, A-285 ASB, Brigham Young University, Provo, UT 84602, irb@byu.edu, (801) 422-1461. El IRB es un grupo de personas que revisa las investigaciones para proteger los derechos y el bienestar de los participantes.

La realización del cuestionario implica su consentimiento en participar. Si decide participar, tenga la bondad de completar el cuestionario adjunto y mándelo por email a jasmine_talbot@byu.edu antes del 18 de octubre. ¡Gracias!
The following consent form in English and Spanish was used in the survey.

Informed Consent Statement

Word Frequency and processing

You have been invited to participate in a research study about the effects of word frequency on the processing of the Spanish language. You were selected as a possible participant because you are currently enrolled in a 300- or 400-level Spanish class at Brigham Young University, or you are a student at BYU and Spanish is your native language. Please read this form and ask any questions you may have before agreeing to participate in the study.

This research study is being conducted by Guinevere Deaver, a Hispanic Linguistics Graduate Student at Brigham Young University in Provo, Utah. Supervising this study is Jeffrey S. Turley, Associate Professor of Spanish and Portuguese at BYU.

Procedure:

If you participate in this study, you will be asked to fill out a 10-minute survey.

Risks and Benefits of Participating in the Study:

There are neither risks nor benefits associated with your participation.

Confidentiality:

The results of the survey will be kept private. No information making it possible to identify you will be included. Research records will be kept in a locked file and only those conducting the research will have access to the records.

Compensation:

There is no compensation for participation in this study.

Voluntary Nature of the Study:

Involvement in this research project is strictly voluntary. You may withdraw at any time without penalty or refuse to participate entirely. There will be no reference to your identity at any point in the research.

Contact Information and Questions:

If you have questions regarding this study, you may contact Guinevere Deaver at handdeaver@byu.edu. You may also contact Dr. Jeffrey Turley at jeffrey_turley@byu.edu. If you have questions regarding your rights as a participant in research projects, you may contact the IRB Administrator, A-285 ASB Campus Drive, Brigham Young University, Provo, UT, 84602; Phone (801) 422-1461; Email irb@byu.edu.

Statement of Consent:

I have read, understood, and received a copy of the above consent and desire to participate in this study of my own free will.

Signature: ____________________________ Date: ____________________________
Declaración de Consentimiento

Procesamiento y frecuencia de palabra

Le invita a usted a participar en una investigación de los efectos de la frecuencia de palabra y del procesamiento mental del español. Usted fue seleccionado como participante porque está matriculado en una clase de español de nivel 300 o 400 de la Universidad Brigham Young, o porque usted es estudiante de BYU y el español es su lengua materna. Tenga la bondad de leer este informe y hacer cualquier pregunta que se le ocurra antes de acceder a participar en el estudio.

Esta investigación está dirigida por Guinevere Deaver, estudiante de posgrado en Lingüística Hispánica de la Universidad Brigham Young en Provo, Utah. El supervisor del estudio es el Dr. Jeffrey S. Turley, Profesor Adjunto de Español y Portugués de BYU.

Método:

Si usted participa en este estudio, tendrá que completar un cuestionario que dura entre 10 y 15 minutos.

Riesgos y beneficios de participar en el estudio:

No hay riesgos ni beneficios asociados con su participación.

Confidencialidad:

Se mantienen privados los resultados del cuestionario. No hay información en el cuestionario que haga posible identificar a los participantes. Las anotaciones relacionadas con el estudio se mantendrán en un archivo guardado bajo llave y solo los que dirigen la investigación tendrán acceso a ellas.

Remuneración:

Ud. no recibirá ninguna remuneración por su participación en el estudio.

Estudio voluntario:

La participación en este estudio es voluntaria. Usted se puede retirar en cualquier momento sin consecuencias o puede negarse a participar totalmente. No habrá ninguna referencia a su identidad en ningún momento del estudio.

Preguntas y datos:

Si Ud. tiene preguntas en cuanto al estudio, sírvase contactar a Guinevere Deaver, handdeaver@byu.edu o al Dr. Jeffrey Turley, Jeffrey_turley@byu.edu. Si tiene alguna pregunta en cuanto a sus derechos como participante de investigación, sírvase contactar al Administrador del IRB, A-285 ASB, Brigham Young University, Provo, UT 84602; irb@byu.edu, (801) 422-1461.

Declaración de consentimiento:

He leído y entendido y recibido una copia del consentimiento y deseo participar por voluntad propia.

Firma: ___________________________________ Fecha: __________________________
Announcement script in English and Spanish for pilot survey

Your input is needed! Guinevere Deaver of the Department of Spanish and Portuguese is researching how Spanish words are processed. She cannot complete her research without help from actual Spanish speakers. Please take 10-15 minutes to complete the following survey after reading the attached consent form. Your response will be anonymous and much appreciated.

Announcement script in English and Spanish for online survey

Your input is needed! Guinevere Deaver of the Department of Spanish and Portuguese is researching how Spanish words are processed. She cannot complete her research without help from actual Spanish speakers. Please take 10-15 minutes to complete the following survey. Please click on the link and read the consent form. A copy of the consent form will be available to print through the survey link. Your response will be anonymous and much appreciated.

Se necesita su ayuda. Guinevere Deaver del Departamento de Español y Portugués está investigando cómo se procesan las palabras en español. Ella no puede completar su investigación sin la ayuda de hispanohablantes. Tenga la bondad de completar la siguiente encuesta de 10 a 15 minutos después de haber leído la declaración de consentimiento. Su participación será muy apreciada.

Se necesita su ayuda. Guinevere Deaver del Departamento de Español y Portugués está investigando cómo se procesan las palabras en español. Ella no puede completar su investigación sin la ayuda de hispanohablantes. Tenga la bondad de completar la siguiente encuesta de 10 a 15 minutos. Haga clic en el enlace y lea la declaración de consentimiento. Se puede imprimir una copia de la declaración de consentimiento que está disponible en la página web de la encuesta. Su participación será muy apreciada.
Appendix C – Pilot Survey Instructions

Preguntas preliminares:

¿Es su lengua materna el español? Sí o No
¿Cursa Ud. una clase de español de nivel 300 o 400? Sí o No

Si usted respondió ‘Sí’ a una de las preguntas, siga leyendo las instrucciones del cuestionario. Si usted respondió ‘No’ a las dos preguntas, no necesita hacer nada más.

Lo siguiente es un experimento con palabras complejas.

Una palabra compleja es una palabra que se puede dividir en unidades más pequeñas que tienen significado.

En español, por ejemplo, la palabra crueldad se divide en dos partes: cruel y dad. –dad es una unidad que aparece al final de muchas palabras españolas. En crueldad, -dad ha sido añadido a cruel para formar la nueva palabra crueldad, que es más compleja. Una palabra que está compuesta por unidades más pequeñas así es una palabra compleja.

Anormal es otro ejemplo de una palabra compleja en español. Se puede separar en dos partes, a- y normal.

Las palabras que no son complejas son simples. Unos ejemplos de palabras simples en español son mesa, amarillo, o cantar. Es imposible dividir la palabra mesa en unidades más pequeñas. Mesa no es compleja.

En el siguiente experimento, se presentarán pares de palabras complejas. Los participantes deben decidir cuál de las dos palabras del par es la MÁS compleja.


Otro ejemplo de una palabra compleja es negritud. Aunque una persona nunca haya escuchado esta palabra antes, podría entender el significado, porque se puede dividir en negro y -tud. Latitud también es compleja – se puede dividir en lato y -tud. Pero latitud no parece tan compleja como negritud. No necesitamos dividir latitud en partes para entender el significado, y, de hecho, no parece natural hacerlo.

Para cada uno de los siguientes pares de palabras, tenga la bondad de leer ambas palabras en voz baja, y luego encierre la palabra más compleja en un círculo. Es muy importante encerrar una palabra de cada par en un círculo, aunque Ud. no esté seguro. No deje ninguna respuesta en blanco. Si es necesario, siga su intuición y haga la mejor selección que pueda.

Ejemplo

[no circulo] negritud [circulo] latitud
Appendix D – Survey Instructions

Cuestionario sobre palabras complejas

Lo siguiente es un experimento con palabras complejas.

Una palabra compleja es una palabra que se puede dividir en unidades más pequeñas que tienen significado.

En español, por ejemplo, la palabra *crueldad* se divide en dos partes: *cruel* y *dad*. *-dad* es una unidad que aparece al final de muchas palabras españolas. En *crueldad*, *-dad* ha sido añadido a *cruel* para formar la nueva palabra *crueldad*, que es más compleja. Una palabra que está compuesta por unidades más pequeñas así es una palabra compleja.

*Contraluz* es otro ejemplo de una palabra compleja en español. Se puede separar en dos partes, *contr-* y *luz*.

Las palabras que no son complejas son simples. Algunos ejemplos de palabras simples en español son *mesa*, *amarillo*, o *arte*. Es imposible dividir la palabra *mesa* en unidades más pequeñas. *Mesa* no es compleja.

En el siguiente experimento, se presentarán pares de palabras complejas. Los participantes deben decidir cuál de las dos palabras del par es la MÁS compleja.


Otro ejemplo de una palabra compleja es *negritud*. Aunque una persona nunca haya escuchado esta palabra antes, podría entender el significado, porque se puede dividir en *negro* y *-tud*. *Gratitud* también es compleja – se puede dividir en *grato* y *-tud*. Pero *gratitud* no parece tan compleja como *negritud*. No necesitamos dividir *gratitud* en partes para entender el significado, y, de hecho, no parece natural hacerlo.

Para cada uno de los siguientes pares de palabras, tenga la bondad de leer ambas palabras en voz baja, y luego seleccione la palabra más compleja, la que es más fácil de dividir en partes que tienen significado. Es muy importante seleccionar una palabra de cada par, aunque Ud. no esté seguro. No deje ninguna respuesta en blanco. Si es necesario, siga su intuición y haga la mejor selección que pueda.

**Ejemplos**

- negritud
- gratitud
- desmontar
- destronar
- sociedad
- sucesión
Appendix E - Survey

The survey as exported from Qualtrics.

Cuestionario sobre palabras complejas

Q1.1 Antes de continuar, tenga la bondad de leer la Declaración de consentimiento. Si decide participar, haga clic en AGREE / ACEPTAR e imprima una copia de la Declaración de consentimiento.

Before continuing, please read the Consent statement. If you decide to participate, click AGREE / ACEPTAR and print a copy of the Consent form.

- AGREE / ACEPTAR (1)
- DISAGREE / NO ACEPTAR (2)

If AGREE / ACEPTAR Is Selected, Then Skip To ¿Es su lengua materna el español?
If DISAGREE / NO ACEPTAR Is Selected, Then Skip To End of Survey

Q1.2 ¿Es su lengua materna el español?

- Sí (1)
- No (2)

If Sí Is Selected, Then Skip To End of Block

Answer If ¿Es su lengua materna el español? No Is Selected

Q1.3 ¿Cursa Ud. una clase de español de nivel 300 o 400?

- Sí (1)
- No (2)

If No Is Selected, Then Skip To End of Survey
Q2.1 ¿Con qué variedades regionales de español ha tenido experiencia?

- el español argentino (1)
- el español boliviano (2)
- el español chileno (3)
- el español colombiano (4)
- el español costarricense (5)
- el español cubano (6)
- el español dominicano (7)
- el español ecuatoriano (8)
- el español de España (9)
- el español guatemalteco (10)
- el español hondureño (11)
- el español mexicano (12)
- el español nicaragüense (13)
- el español panameño (14)
- el español paraguayo (15)
- el español peruano (16)
- el español puertorriqueño (17)
- el español salvadoreño (18)
- el español uruguayo (19)
- el español venezolano (20)
- otra variedad (21) ____________________

Q3.1 ¿Cuánto tiempo estuvo en contacto con ${lm://Field/1}$?

Q4.1 ¿Ha pasado tiempo en un país de habla hispana?

- Sí (1)
- No (2)
Answer If ¿Ha pasado tiempo en un país de habla hispana? Sí is Selected

Q4.2 ¿En cuáles países?
- Argentina (1)
- Bolivia (2)
- Chile (3)
- Colombia (4)
- Costa Rica (5)
- Cuba (6)
- Ecuador (7)
- El Salvador (8)
- España (9)
- Guatemala (10)
- Honduras (11)
- México (12)
- Nicaragua (13)
- Panamá (14)
- Paraguay (15)
- Perú (16)
- Puerto Rico (17)
- República Dominicana (18)
- Uruguay (19)
- Venezuela (20)

Q5.1 ¿Cuánto tiempo estuvo en ${lm://Field/1}$?
- 3 meses o menos (1)
- 4-6 meses (2)
- 7-11 meses (3)
- un año (4)
- un año y medio (5)
- 2 años (6)
- 2 años y medio (7)
- 3 años (8)
- más de 3 años (9)

Q6.1 Si entiende las instrucciones, seleccione ENTENDIDO y siga adelante. Si no entiende las instrucciones, léalas otra vez.
- ENTENDIDO (1)
Q7.1 ¿Cuál variedad regional de español habla Ud.?
- el español argentino (1)
- el español boliviano (2)
- el español chileno (3)
- el español colombiano (4)
- el español costarricense (5)
- el español cubano (6)
- el español dominicano (7)
- el español ecuatoriano (8)
- el español de España (9)
- el español guatemalteco (10)
- el español hondureño (11)
- el español mexicano (12)
- el español nicaragüense (13)
- el español panameño (14)
- el español paraguayo (15)
- el español peruano (16)
- el español puertorriqueño (17)
- el español salvadoreño (18)
- el español uruguayo (19)
- el español venezolano (20)

Q8.1 ¿Ha recibido educación formal en ${lm://Field/1}?
- Sí (1)
- No (2)

Answer If ¿Ha recibido educación formal en ${lm://Field/1} (Any Loop) Sí Is Selected
Q8.2 ¿Cuántos años de educación formal recibió Ud. en ${lm://Field/1}?

Q8.3 ¿Por cuántos años ha hablado ${lm://Field/1}?

Q9.1 Seleccione la palabra más compleja.
- carpintero (1)
- carnicero (2)

Q9.2 Seleccione la palabra más compleja.
- amplificador (1)
- acelerador (2)

Q9.3 Seleccione la palabra más compleja.
- ritual (1)
- visual (2)
Q9.4 Seleccione la palabra más compleja.
- desdénar (1)
- desconfiar (2)

Q9.5 Seleccione la palabra más compleja.
- avidez (1)
- acidez (2)

Q9.6 Seleccione la palabra más compleja.
- posición (1)
- perdición (2)

Q9.7 Seleccione la palabra más compleja.
- inmutable (1)
- inestable (2)

Q9.8 Seleccione la palabra más compleja.
- soñador (1)
- orador (2)

Q9.9 Seleccione la palabra más compleja.
- sobrecogedor (1)
- sobrenatural (2)

Q9.10 Seleccione la palabra más compleja.
- novedad (1)
- voluntad (2)

Q9.11 Seleccione la palabra más compleja.
- observancia (1)
- relevancia (2)

Q9.12 Seleccione la palabra más compleja.
- irresoluto (1)
- irrepetido (2)

Q9.13 Seleccione la palabra más compleja.
- desencanto (1)
- desayuno (2)

Q9.14 Seleccione la palabra más compleja.
- prejuzgar (1)
- preservar (2)
Q9.15 Seleccione la palabra más compleja.
- despejar (1)
- deschar (2)

Q9.16 Seleccione la palabra más compleja.
- objetividad (1)
- creatividad (2)

Q9.17 Seleccione la palabra más compleja.
- discontinuo (1)
- disyuntivo (2)

Q9.18 Seleccione la palabra más compleja.
- asidero (1)
- hervidero (2)

Q9.19 Seleccione la palabra más compleja.
- infinito (1)
- indirecto (2)

Q9.20 Seleccione la palabra más compleja.
- estadístico (1)
- estilístico (2)

Q9.21 Seleccione la palabra más compleja.
- previsión (1)
- prevención (2)

Q9.22 Seleccione la palabra más compleja.
- honestad (1)
- majestad (2)

Q9.23 Seleccione la palabra más compleja.
- fundador (1)
- fumador (2)

Q9.24 Seleccione la palabra más compleja.
- consiguiente (1)
- inteligente (2)

Q9.25 Seleccione la palabra más compleja.
- actividad (1)
- humanidad (2)
Q9.26 Seleccione la palabra más compleja.
- inaceptable (1)
- inagotable (2)

Q9.27 Seleccione la palabra más compleja.
- subrayar (1)
- subvenir (2)

Q9.28 Seleccione la palabra más compleja.
- paciente (1)
- creciente (2)

Q9.29 Seleccione la palabra más compleja.
- petición (1)
- edición (2)

Q9.30 Seleccione la palabra más compleja.
- intercelular (1)
- interlocutor (2)

Q9.31 Seleccione la palabra más compleja.
- ambición (1)
- medición (2)

Q9.32 Seleccione la palabra más compleja.
- intangible (1)
- inaudible (2)

Q9.33 Seleccione la palabra más compleja.
- ambigüedad (1)
- antigüedad (2)

Q9.34 Seleccione la palabra más compleja.
- transformar (1)
- transportar (2)

Q9.35 Seleccione la palabra más compleja.
- oriente (1)
- sonriente (2)

Q9.36 Seleccione la palabra más compleja.
- pretender (1)
- preceder (2)
Q9.37 Seleccione la palabra más compleja.
- anticipo (1)
- anticuerpo (2)

Q9.38 Seleccione la palabra más compleja.
- mortífero (1)
- mamífero (2)

Q9.39 Seleccione la palabra más compleja.
- nadador (1)
- labrador (2)

Q9.40 Seleccione la palabra más compleja.
- increíble (1)
- invisible (2)

Q9.41 Seleccione la palabra más compleja.
- plenitud (1)
- magnitud (2)

Q9.42 Seleccione la palabra más compleja.
- superego (1)
- superficie (2)

Q9.43 Seleccione la palabra más compleja.
- antecesor (1)
- anteponer (2)

Q9.44 Seleccione la palabra más compleja.
- preconocer (1)
- prevalecer (2)

Q9.45 Seleccione la palabra más compleja.
- irrelevante (1)
- irreverente (2)

Q9.46 Seleccione la palabra más compleja.
- sobrevenida (1)
- sobresaliente (2)

Q9.47 Seleccione la palabra más compleja.
- contradictor (1)
- contrapesar (2)
Q9.48 Seleccione la palabra más compleja.
- partición (1)
- nutrición (2)

Q9.49 Seleccione la palabra más compleja.
- frecuencia (1)
- secuencia (2)

Q9.50 Seleccione la palabra más compleja.
- imperio (1)
- impuro (2)

Q9.51 Seleccione la palabra más compleja.
- irresponsable (1)
- irreprochable (2)

Q9.52 Seleccione la palabra más compleja.
- retroproyector (1)
- retrospectiva (2)

Q9.53 Seleccione la palabra más compleja.
- delatar (1)
- desatar (2)

Q9.54 Seleccione la palabra más compleja.
- repetición (1)
- superstición (2)

Q9.55 Seleccione la palabra más compleja.
- habilidad (1)
- debilidad (2)

Q9.56 Seleccione la palabra más compleja.
- subvenir (1)
- sublevar (2)

Q9.57 Seleccione la palabra más compleja.
- virtud (1)
- quietud (2)

Q9.58 Seleccione la palabra más compleja.
- tradicional (1)
- profesional (2)
Q9.59 Seleccione la palabra más compleja.
- disconforme (1)
- disparate (2)

Q9.60 Seleccione la palabra más compleja.
- serpiente (1)
- valiente (2)
## Appendix F – Survey Responses by Item

| Survey Item no. | Word A | Surface Freq. | Base Freq. | Token Total | Word B | Surface Freq. | Base Freq. | Token Total | Native Speaker Responses by Item | | Native Speaker Responses by Item | | Native Speaker Responses by Item | \hline
| 2 | amplificador | 5 | 1 | 8 | acelerador | 7 | 0 | 16 | | | | | \hline
| 5 | asidero | 23 | 13 | 36 | asidero | 14 | 55 | 79 | | | | | \hline
| 7 | inmutable | 23 | 1 | 24 | inmutable | 11 | 35 | 46 | | | | | \hline
| 8 | orador | 6 | 2 | 8 | orador | 9 | 51 | 40 | | | | | \hline
| 9 | sobrecogedor | 7 | 0 | 7 | sobrecogedor | 17 | 334 | 551 | | | | | \hline
| 11 | relevancia | 8 | 1 | 9 | observancia | 8 | 82 | 85 | | | | | \hline
| 12 | irresoluto | 2 | 1 | 3 | irresoluto | 1 | 29 | 30 | | | | | \hline
| 13 | desayuno | 62 | 2 | 64 | desayuno | 16 | 44 | 60 | | | | | \hline
| 16 | creatividad | 22 | 11 | 33 | creatividad | 20 | 108 | 128 | | | | | \hline
| 18 | asidero | 5 | 2 | 7 | asidero | 3 | 4 | 7 | | | | | \hline
| 19 | infinito | 73 | 16 | 89 | indirecto | 7 | 53 | 60 | | | | | \hline
| 23 | fundador | 17 | 13 | 30 | fundador | 10 | 65 | 73 | | | | | \hline
| 25 | actividad | 240 | 56 | 296 | humanidad | 121 | 240 | 361 | | | | | \hline
| 26 | inmutable | 13 | 1 | 14 | inmutable | 6 | 18 | 24 | | | | | \hline
| 27 | subrayar | 12 | 1 | 13 | subrayar | 1 | 123 | 124 | | | | | \hline
| 29 | edición | 53 | 4 | 57 | edición | 53 | 99 | 152 | | | | | \hline
| 30 | interlocutor | 33 | 30 | 63 | interlocutor | 2 | 89 | 91 | | | | | \hline
| 52 | manifiesto | 4 | 4 | 8 | manifiesto | 4 | 11 | 15 | | | | | \hline
| 55 | embriagado | 46 | 51 | 77 | embriagado | 45 | 174 | 219 | | | | | \hline
| 34 | transportar | 11 | 3 | 14 | transformar | 5 | 57 | 88 | | | | | \hline
| 56 | pretender | 25 | 8 | 33 | pretender | 2 | 24 | 25 | | | | | \hline
| 39 | Embargo | 9 | 1 | 10 | embargo | 6 | 20 | 26 | | | | | \hline
| 40 | increible | 50 | 4 | 54 | increible | 54 | 57 | 111 | | | | | \hline
| 41 | magnitud | 43 | 8 | 51 | magnitud | 58 | 92 | 150 | | | | | \hline
| 45 | irreverente | 4 | 3 | 7 | irrelevante | 8 | 11 | 19 | | | | | \hline
| 46 | sobresaliente | 9 | 3 | 12 | sobrevivencia | 2 | 12 | 14 | | | | | \hline
| 48 | nutrición | 9 | 1 | 10 | nutrición | 7 | 342 | 549 | | | | | \hline
| 51 | irrespetable | 6 | 1 | 7 | irrespetable | 9 | 74 | 83 | | | | | \hline
| 55 | habilidad | 52 | 15 | 67 | habilidad | 41 | 68 | 109 | | | | | \hline
| 58 | profesional | 103 | 69 | 172 | profesional | 69 | 151 | 220 | | | | | \hline
<table>
<thead>
<tr>
<th>Survey Item no.</th>
<th>Pseudo affixed</th>
<th>Freq.</th>
<th>Affixed word</th>
<th>Freq.</th>
<th>Filler</th>
<th>Affixed</th>
<th>Filler</th>
<th>Affixed</th>
<th>Filler</th>
<th>Affixed</th>
<th>2 Native Speakers In Spanish Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>carpintero</td>
<td>12</td>
<td>carpintero</td>
<td>6</td>
<td>8</td>
<td>3</td>
<td>4</td>
<td>22</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>visual</td>
<td>38</td>
<td>ritual</td>
<td>53</td>
<td>6</td>
<td>5</td>
<td>12</td>
<td>14</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>desedar</td>
<td>7</td>
<td>desconfiar</td>
<td>8</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>23</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>posicion</td>
<td>166</td>
<td>perdicion</td>
<td>18</td>
<td>2</td>
<td>9</td>
<td>2</td>
<td>24</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>voluntad</td>
<td>234</td>
<td>movilidad</td>
<td>62</td>
<td>3</td>
<td>8</td>
<td>4</td>
<td>22</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>preservar</td>
<td>20</td>
<td>preservar</td>
<td>15</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>25</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>despellar</td>
<td>7</td>
<td>desear</td>
<td>8</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>22</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td>disuasivo</td>
<td>1</td>
<td>discontinuo</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>22</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>estadistica</td>
<td>9</td>
<td>estadística</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>19</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>21</td>
<td>prevencion</td>
<td>19</td>
<td>prevencion</td>
<td>7</td>
<td>3</td>
<td>8</td>
<td>5</td>
<td>21</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>majestad</td>
<td>60</td>
<td>honestad</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>8</td>
<td>18</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>24</td>
<td>inteligente</td>
<td>92</td>
<td>consiguiente</td>
<td>60</td>
<td>1</td>
<td>10</td>
<td>2</td>
<td>24</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>28</td>
<td>paciente</td>
<td>52</td>
<td>creciente</td>
<td>63</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>23</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>31</td>
<td>ambicion</td>
<td>27</td>
<td>medicion</td>
<td>15</td>
<td>5</td>
<td>6</td>
<td>13</td>
<td>13</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>35</td>
<td>oriente</td>
<td>47</td>
<td>sonriente</td>
<td>48</td>
<td>4</td>
<td>7</td>
<td>2</td>
<td>24</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>37</td>
<td>anticipo</td>
<td>7</td>
<td>antecuerpo</td>
<td>16</td>
<td>3</td>
<td>8</td>
<td>4</td>
<td>22</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>38</td>
<td>manifiesto</td>
<td>6</td>
<td>manifiesto</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>13</td>
<td>11</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>42</td>
<td>superficie</td>
<td>162</td>
<td>superficie</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>3</td>
<td>23</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>43</td>
<td>antecesor</td>
<td>9</td>
<td>antecesor</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>2</td>
<td>24</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>44</td>
<td>prevaler</td>
<td>5</td>
<td>prevaler</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>24</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>47</td>
<td>contradecir</td>
<td>2</td>
<td>contradecir</td>
<td>2</td>
<td>3</td>
<td>8</td>
<td>8</td>
<td>18</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>49</td>
<td>secuencia</td>
<td>55</td>
<td>frecuencia</td>
<td>234</td>
<td>5</td>
<td>6</td>
<td>9</td>
<td>17</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>50</td>
<td>imperio</td>
<td>98</td>
<td>imperio</td>
<td>5</td>
<td>2</td>
<td>9</td>
<td>4</td>
<td>22</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>52</td>
<td>retrospectiva</td>
<td>7</td>
<td>retrospectiva</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>8</td>
<td>18</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>53</td>
<td>delatar</td>
<td>4</td>
<td>delatar</td>
<td>7</td>
<td>1</td>
<td>10</td>
<td>8</td>
<td>18</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>54</td>
<td>supersticion</td>
<td>12</td>
<td>repeticion</td>
<td>36</td>
<td>6</td>
<td>5</td>
<td>12</td>
<td>14</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>55</td>
<td>sublevar</td>
<td>1</td>
<td>subvencion</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>22</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>57</td>
<td>virtud</td>
<td>95</td>
<td>quietud</td>
<td>44</td>
<td>3</td>
<td>8</td>
<td>4</td>
<td>22</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>59</td>
<td>disparate</td>
<td>12</td>
<td>desacordo</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>1</td>
<td>25</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>60</td>
<td>serpent fee</td>
<td>35</td>
<td>valiente</td>
<td>31</td>
<td>4</td>
<td>7</td>
<td>8</td>
<td>18</td>
<td></td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>
## Appendix G – Survey Responses by Participant

### Responses by Participant

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Native Speaker Participant</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word A</td>
<td>17</td>
<td>11</td>
<td>17</td>
<td>12</td>
<td>14</td>
<td>17</td>
<td>15</td>
<td>16</td>
<td>14</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Word B</td>
<td>13</td>
<td>19</td>
<td>13</td>
<td>18</td>
<td>16</td>
<td>13</td>
<td>15</td>
<td>14</td>
<td>16</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Filler</td>
<td>10</td>
<td>2</td>
<td>16</td>
<td>5</td>
<td>17</td>
<td>14</td>
<td>14</td>
<td>13</td>
<td>10</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td><strong>L2 Spanish Speaker Participant</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word A</td>
<td>14</td>
<td>16</td>
<td>13</td>
<td>12</td>
<td>10</td>
<td>8</td>
<td>9</td>
<td>16</td>
<td>14</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>Word B</td>
<td>16</td>
<td>14</td>
<td>17</td>
<td>18</td>
<td>20</td>
<td>22</td>
<td>21</td>
<td>14</td>
<td>14</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td>Filler</td>
<td>10</td>
<td>12</td>
<td>5</td>
<td>6</td>
<td>11</td>
<td>5</td>
<td>1</td>
<td>15</td>
<td>14</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td><strong>L2 Spanish Speaker cont.</strong></td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
</tr>
<tr>
<td>Word A</td>
<td>13</td>
<td>14</td>
<td>6</td>
<td>9</td>
<td>9</td>
<td>8</td>
<td>16</td>
<td>9</td>
<td>6</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>Word B</td>
<td>17</td>
<td>16</td>
<td>24</td>
<td>21</td>
<td>21</td>
<td>22</td>
<td>14</td>
<td>21</td>
<td>24</td>
<td>21</td>
<td>18</td>
</tr>
<tr>
<td>Filler</td>
<td>12</td>
<td>9</td>
<td>4</td>
<td>8</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td><strong>Native Participant in Spanish Class</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word A</td>
<td>6</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word B</td>
<td>24</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Filler</td>
<td>3</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Bibliography


