

# Orthographic Similarity and False Recognition for Unfamiliar Words

2015

Jeffrey Perrotte  
*University of Central Florida*

Find similar works at: <https://stars.library.ucf.edu/honorstheses1990-2015>

University of Central Florida Libraries <http://library.ucf.edu>

 Part of the [Psychology Commons](#)

## Recommended Citation

Perrotte, Jeffrey, "Orthographic Similarity and False Recognition for Unfamiliar Words" (2015). *HIM 1990-2015*. 1879.  
<https://stars.library.ucf.edu/honorstheses1990-2015/1879>

This Open Access is brought to you for free and open access by STARS. It has been accepted for inclusion in HIM 1990-2015 by an authorized administrator of STARS. For more information, please contact [lee.dotson@ucf.edu](mailto:lee.dotson@ucf.edu).

ORTHOGRAPHIC SIMILARITY AND FALSE RECOGNITION FOR  
UNFAMILIAR WORDS

by

JEFFREY A. PERROTTE

A thesis submitted in partial fulfillment of the requirements  
for the Honors in the Major Program in Psychology  
in the College of Sciences  
and in the Burnett Honors College  
at the University of Central Florida  
Orlando, Florida

Fall Term, 2015

Thesis Chair: Marisol Parra-Tatge, Ph.D.

## ABSTRACT

There is evidence of false recognition (FR) driven by orthographic similarities within languages (Lambert, Chang, & Lin, 2001; Raser, 1972) and some evidence that FR crosses languages (Parra, 2013). No study has investigated whether FR based on orthographic similarities occurs for unknown words in an unknown language. This study aimed to answer this question. It further explored whether FR based on orthographic similarities is more likely in a known (English) than in an unknown (Spanish) language. Forty-six English monolinguals participated. They studied 50 English and 50 Spanish words during a study phase. A recognition test was given immediately after the study phase. It consisted of 40 Spanish and 40 English words. It included list words (i.e., words presented at study); homographs (i.e., words not presented at study, orthographically similar to words presented at study); and unrelated words (i.e., words not presented at study, not orthographically similar to words presented at study). The LSD post-hoc test showed significant results supporting the hypothesis that false recognition based on orthographic similarities occurs for words in a known language (English) and in an unknown language (Spanish). Further evidence was provided by the LSD post-hoc test supporting the hypothesis that false recognition based on orthographic similarities was more likely to occur in a known language than an unknown language. Results provided evidence that the meaning and orthographic form are used when information is encoded thereby influencing recognition decisions. Furthermore, these results emphasize the significance of orthography when information is encoded and retrieved.

*Keywords:* false recognition, orthography, semantic, orthographic distinctiveness, semantic distinctiveness, English monolingual, least significant difference (LSD)

## **DEDICATION**

For my daughter, she has inspired me  
to better myself everyday.

For my mother, she has given me her  
support and wisdom to continue on.

## **ACKNOWLEDGEMENTS**

I would like to thank my mentor, Dr. Marisol Parra-Tatge for her advice, guidance, and dedication throughout this entire thesis process.

## TABLE OF CONTENTS

<b>INTRODUCTION.....</b>	<b>1</b>
<b>Semantic Similarity and False Recognition in the Native Language .....</b>	<b>1</b>
<b>Orthographic Similarity and False Recognition in the Native Language .....</b>	<b>2</b>
<b>Semantic Similarity and False Recognition across Languages.....</b>	<b>3</b>
<b>Orthographic Similarity and False Recognition across Languages .....</b>	<b>4</b>
<b>Purpose and Hypotheses.....</b>	<b>5</b>
Word frequency effect .....	5
Orthographic distinctiveness hypothesis .....	6
Semantic distinctiveness hypothesis .....	6
<b>Expected Results .....</b>	<b>7</b>
<b>METHOD .....</b>	<b>9</b>
<b>Participants.....</b>	<b>9</b>
<b>Materials .....</b>	<b>10</b>
Language Questionnaire .....	10
Study words .....	11
Recognition test words.....	11
<b>Procedure.....</b>	<b>12</b>
<b>RESULTS .....</b>	<b>14</b>
<b>False Recognition for known and unknown words similar in orthography .....</b>	<b>14</b>
<b>Comparison of word types for known and unknown words.....</b>	<b>15</b>
Homographs .....	16
List words .....	16
Unrelated words .....	17
<b>DISCUSSION .....</b>	<b>18</b>
<b>APPENDIX A: LANGUAGE QUESTIONNAIRE .....</b>	<b>22</b>

<b>APPENDIX B: STUDY WORD BOOKLET .....</b>	<b>26</b>
<b>APPENDIX C: RECOGNITION TEST BOOKLET.....</b>	<b>39</b>
<b>APPENDIXD: IRB APROVAL LETTER .....</b>	<b>44</b>
<b>REFERENCES.....</b>	<b>47</b>

## **INTRODUCTION**

In the past, studies regarding word recognition placed much emphasis on the effects of word meaning on word recognition, particularly false recognition (Cabeza & Lennartson, 2005; Collins & Loftus, 1975; Roediger & McDermott, 1995). However, in recent years there has been an increasing interest in the role of orthographic similarity (e.g., Hunt & Elliot, 1980; Lambert, Chang, & Lin, 2001). There is evidence that orthographic similarity influences word recognition on the native language (Hunt & Elliott, 1980; Lambert, Chang, & Lin, 2001) and across languages (Parra, 2013). The present study will examine whether false recognition based on orthographic similarities also occurs for unknown words in an unknown language.

### **Semantic Similarity and False Recognition in the Native Language**

The Deese-Roediger-McDermott (Deese, 1959; Roediger & McDermott, 1995), DRM paradigm, has been used in previous studies to examine whether false recognition occurs for words similar in meaning (i.e., semantic similarity) in the native language. The DRM paradigm consists of two phases, the study phase and the recognition test. At the study phase, words are presented that are similar in meaning to one another (e.g., awake, bed, rest, tired, dream, wake) and similar in meaning to words that will be presented at recognition (e.g., sleep), known as the critical lure. The recognition test consists of words that were presented at study (i.e., list words), words that were not presented at study (i.e., unrelated words), and the critical lure. Using a list of words semantically associated to one another (i.e., DRM list), Roediger and McDermott (1995) showed that the critical lure was recognized as well or better than the unrelated list words. According to Roediger and McDermott, false recognition of the critical lure occurred because the



words similar in meaning presented at study activated other words in the semantic network including the critical lure. The critical lure would be strongly activated because it would receive the combined activation of the words in a DRM list. Thus, the critical lure will be more likely to be falsely recognized than other non-presented words that were not activated at study.

### **Orthographic Similarity and False Recognition in the Native Language**

In addition to false recognition occurring for words similar in meaning, false recognition has been shown to occur for words similar in form (i.e., orthographic form). Words that were not presented at study but that are orthographically similar to words presented at study are more likely to be falsely recognized than words that were not presented at study and were not orthographically similar to words presented at study (Lambert, Chang, & Lin, 2001; Raser, 1972). This was demonstrated by Raser (1972) using a recognition task to show how false recognition occurs for words similar in orthography. Participants were presented with a list of words and then were given a recognition test that consisted of words previously presented at the study phase (i.e., list words), words that were not presented at study but were either high or low in orthographic similarity to words presented at study, and words not presented at study and not similar in orthography to words presented at study (i.e., unrelated words). Words that were defined as high in orthographic similarity were words that had the same length (i.e., same amount of letters) and were different by only one letter other than the first or last letter (e.g., *sour* and *slur*) relative to list words. Words that were considered low in orthographic similarity were words that differed by two or more letters and were not the same length (e.g., *cloud* and *cleave*). Raser concluded that words high in orthographic similarity were more likely to be falsely recognized than words low in orthographic similarity and unrelated words. Additional evidence

has been reported that false recognition based on orthographic similarities occurs for words in the native language. A recognition task was conducted by Lambert, Chang, and Lin (2001) that examined false recognition for words similar in orthography. Participants were presented with words at study and then given a recognition test that consisted of words similar in orthographic form to presented words, list words, and unrelated words. Orthographic similarity of the words at recognition were defined by the amount of two-letter sub-sequences shared between two words, known as bigrams. For example, bigrams for the words *alive* {\_a, al, li, iv, ve, e\_} and *alike* {\_a, al, li, ik, ke, e\_}, share a total of 4 bigrams {\_a, al, li, e\_}. Lambert et al. concluded that the probability of false recognition occurring for words similar in orthographic form to presented words increases as the number of orthographic similarities between words increases (i.e., the amount of bigrams words share).

### **Semantic Similarity and False Recognition across Languages**

False recognition has not only been shown to occur for orthographic and semantic similarities for words in the native language, but it has also been demonstrated to occur across languages. An adapted version of the DRM paradigm was used by Cabeza and Lennartson (2005) that examined false recognition occurring based on semantic similarities in English-French bilinguals. Cabeza and Lennartson adapted the DRM paradigm by including English and French words in different study-test conditions (i.e., English-English, French-French, English-French, and French-English) to investigate within and cross-language influences on false recognition in bilinguals. In the study phase, bilinguals were presented with DRM lists of words in either English or French. The recognition test consisted of English and French list words, unrelated words, and critical lures. The findings of this study indicate that false recognition was

more likely to occur for the critical lure than for unrelated words and just as likely to occur for list words when words presented at study were in one language and words at recognition were in the other language. The authors concluded that the high occurrence of false recognition for critical lures in bilinguals was the result of shared semantic representations of words in both languages. Words activated at study in one language share the same meaning and represent the same concept in the other language. In this context, when the DRM list words are studied in one language, words semantically related to the list words (e.g., critical lure) are activated in both languages in the semantic network.

### **Orthographic Similarity and False Recognition across Languages**

There is some evidence that false recognition based in orthographic similarities occurs across languages in bilinguals. Parra (2013) demonstrated that false recognition occurs across languages in English-Spanish bilinguals for words similar in orthographic form using the modified version of the DRM paradigm. Participants were presented with English and Spanish DRM lists during the study phase. The recognition test consisted of English and Spanish words, as follows: non-presented words that were orthographically similar to presented words (i.e., interlingual homographs), list words, non-presented words similar in meaning to presented words (i.e., critical lures), non-presented words that were similar in form and meaning (i.e., cognates), and non-presented words that were neither similar in meaning nor similar in form to presented words. The orthographic similarity of the interlingual homographs were defined by a difference in form of 1-3 letters and a difference in length by 1 or 2 letters. Parra reported marginally significant results that false recognition occurred for words similar in orthographic form to words previously presented at study across languages in bilinguals. She concluded that the initial

activation of words in one language at study, activate other non-studied words that are similar in orthographic form in both languages for bilinguals. This indicates that although the lexica of the two languages are separate, they interact.

### **Purpose and Hypotheses**

False recognition of non-presented words similar in orthography to presented words in an unfamiliar language has not been investigated. One of the purposes of this study was to investigate whether false recognition based on orthographic similarities occurs for unknown words of an unfamiliar language in English monolinguals.

This study further examined whether false recognition is more likely to occur in the native language, English, than in an unfamiliar language, Spanish. Different predictions were made according to the word frequency effect, and the orthographic and semantic distinctiveness hypotheses.

#### ***Word frequency effect***

According to the word frequency effect, low frequency words are better recognized than high frequency words. Low frequency words are words used less often in a language (e.g., *naive*) and high frequency words are words used most often in everyday language (e.g., *the*). Words in an unfamiliar language are considered extremely low frequency words. Consistent with the word frequency effect, they would be better recognized than words in a familiar language (which for native speakers of that language are high frequency words). In terms of false recognition, for monolinguals, words in an unknown language would be less likely to be falsely recognized whereas words in the native language would more likely be falsely recognized (because words of

an unfamiliar language are more distinct and then would be better remembered than words of a familiar language).

### ***Orthographic distinctiveness hypothesis***

In line with the word frequency effect is the orthographic distinctiveness hypothesis. According to this hypothesis, orthographically distinctive words are better recognized than words that are not orthographically distinct (Hunt & Elliott, 1980). Hunt and Elliott describe orthographic distinctive words, as words that have uncommon structures or features (e.g., phlegm). Different languages have unique combinations of sounds and letters. Thus, they have distinctive features for people who are not familiar with a particular language. Since they are more distinct than words in the native language, words in an unfamiliar language would be remembered better and, then, would be less likely to be falsely recognized than words in the native language (words in a familiar language).

### ***Semantic distinctiveness hypothesis***

An alternative to the word frequency effect and the orthographic distinctiveness hypothesis is the semantic distinctiveness hypothesis (Ozubko & Joordens, 2011). According to this hypothesis, the meaning of words can help differentiate words similar in orthographic form (e.g., *horse* vs. *house*). Monolinguals know the meaning of words in their native language and this knowledge will help them discriminate between words that are orthographically similar. Thus, in this view, false recognition would be more likely to occur for words in an unfamiliar language (individuals do not know the meaning of words in that language, and would be less likely to use meaning to differentiate between two orthographically similar words) than for

words in the native language. Regarding list words, words in a familiar language would be better recognized than words in an unfamiliar language.

## **Expected Results**

In the present study, participants were English native monolinguals. A recognition task was used that consisted of two phases, the study phase and the recognition test. In the study phase, English monolinguals were presented with words in English and in Spanish (see Appendix B). Then a recognition test was given to participants that consisted of words previously presented at study (i.e., list words), words that were not presented at study (i.e., unrelated words), and words not presented at study that are similar in orthographic form to presented words (i.e., homographs; see Appendix C). Participants selected either yes (indicating that they recognize the word from the study phase) or they selected no (indicating that they did not recognize the words from the study phase) for each word at the recognition phase.

Consistent with the word frequency effect and the orthographic distinctiveness hypothesis, it was expected that Spanish list words (that are less frequent and more distinctive than English list words) would be better recognized than English list words. According to the semantic distinctiveness hypothesis, the opposite would be true. English list words would be better recognized than Spanish list words.

Regarding false recognition, according to the orthographic distinctiveness hypothesis, English homographs would be more falsely recognized than Spanish homographs (assuming that participants would remember better Spanish words because they are more orthographically distinct). In contrast, consistent with the semantic distinctiveness hypothesis, Spanish

homographs would be more falsely recognized than English homographs (assuming that participants would resort to the meaning of English words to differentiate between orthographically similar English words).

## METHOD

### Participants

The total participants were forty-six ( $n = 46$ ) University of Central Florida college students comprised of 14 males (30.4%) and 32 females (69.6%). The age of participants ranged from 18 to 23 years old ( $M = 18.39$ ,  $SD = 1.0$ ). Out of the 46 participants, 37 were Freshmen (80.4%), 6 were Sophomores (12.8%), 2 were Juniors (4.3%), and 1 was a Senior (2.2%). All participants were English speaking monolinguals who cannot read, speak, or understand a second language. Participants had no history of hearing impairments, language disabilities, or learning disabilities. Participants had a basic knowledge of the Spanish language. A total of 34 participants (74.5%) had taken Spanish courses in High School. Of those 34 participants, 1 had taken Spanish for one year (3.0%), 23 had taken Spanish for two years (69.7%), and 9 had taken Spanish for three years (27.3%). However, self-ratings of their Spanish proficiency provided by participants showed that their proficiency in Spanish was low. Participants provided self-proficiency ratings regarding speaking Spanish, understanding spoken Spanish, and reading Spanish. Self-proficiency ratings were recorded using a 10-point Likert scale where 0 indicated *no proficiency* and 10 indicated *extremely proficient* (see Appendix A). Proficiency ratings in speaking Spanish ranged from 0 to 3 ( $M = 0.83$ ,  $SD = 0.94$ ); proficiency ratings in understanding spoken Spanish ranged from 0 to 3 ( $M = 1.10$ ,  $SD = 0.95$ ); and proficiency ratings in reading Spanish ranged from 0 to 3 ( $M = 1.10$ ,  $SD = 1.12$ ). Participants also provided self-ratings on the amount of exposure to the Spanish language in several categories. Exposure self-ratings were recorded using a 10-point Likert scale where 0 indicated *no exposure* and 10 indicated an *extreme amount of exposure* (see Appendix A). Out of 46 participants, only 3 participants (6.4%)



reported exposure to the Spanish language. In addition, self-ratings provided by the three participants regarding the amount of exposure to the Spanish language were minimal. Regarding the three participants who reported exposure to the Spanish language, the amount of exposure to the Spanish language from interacting with friends ranged from 0 to 3 ( $M = 0.15$ ,  $SD = 0.63$ ); exposure from interacting with family ranged from 0 to 1 ( $M = 0.02$ ,  $SD = 0.15$ ); exposure from watching television ranged from 0 to 2 ( $M = 0.07$ ,  $SD = 0.33$ ); exposure from listening to the radio or music ranged from 0 to 2 ( $M = 0.11$ ,  $SD = 0.43$ ); exposure from work ranged from 0 to 1 ( $M = 0.02$ ,  $SD = 0.15$ ); and no participants reported exposure to the Spanish language by reading Spanish or through self-instruction or language lab.

## **Materials**

### ***Language Questionnaire***

A modified version of The Language Experience and Proficiency Questionnaire (LEAP-Q) (Marian, Blumenfeld, & Kaushanskaya, 2007) was given to participants in the experimental part of the study and consisted of 14 questions. The purpose of the Language Questionnaire was to collect demographical data such as race, gender, age, and it included questions regarding language history of the participants. For example, “Can you understand a spoken language other than your native language?” Other questions pertained to the level of exposure to the Spanish language in seven different categories using a 10-point Likert scale. For example, Rate your exposure to the Spanish language from watching television with 0 being the *lowest* amount of exposure and 10 being the *highest* amount of exposure (see Appendix A).

### ***Study words***

The participants were given a booklet composed of 100 words, 50 of these words were English words and 50 were Spanish words. The English words were taken from the *Brown Corpus* (Kucera & Francis, 1967) that is comprised of over one million words. The 50 English words extracted from the *Brown Corpus* had a total mean frequency of 5,943.02 ( $SD = 8,829.73$ ). The English words were one or two syllables with a minimum of 4 letters and a maximum 7 letters. The 50 Spanish words were taken from an online corpus CREA (Corpus de Referencia del Español Actual) (Spanish Royal Academy) and had a total mean frequency of  $M = 4,764.42$  ( $SD = 10,083.69$ ). The Spanish words were also two or three syllables with a minimum of 4 letters and a maximum of 7 letters. The English and Spanish words were randomized on 10 separate pages consisting of 10 words on each page. The study words were typed on 8.5X11 printer paper and centered with 16' Times New Roman black font (see Appendix B).

### ***Recognition test words***

Participants were given a booklet of 80 randomized words, 40 English words and 40 Spanish words. The English words were one or two syllables with a minimum of 4 letters and a maximum of 7 letters. The Spanish words were two or three syllables with a minimum of 4 letters and a maximum of 7 letters. For each word there was a “No” or “Yes” response box. Participants marked the “No” response box if they *did not* recognize the word as the same word presented in the study phase of the experiment (i.e., study word). Participants marked the “Yes” response box if they *did* recognize the word as the same word presented in the study phase of the

experiment. There was a total of 80 recognition test words, 40 of these words were the same words that were presented in the study phase of the experiment (i.e., list words) and consisted of 20 English words and 20 Spanish words. Of the 40 remaining recognition test words, 20 of these words were orthographically similar (i.e., homographs) to the words presented in the study phase of the experiment and 10 of the homographs were English words and 10 of the homographs were Spanish words; And 20 of the words were words that were not presented (i.e., unrelated) in the study phase of the experiment and consisted of 10 English words and 10 Spanish words. The 20 unrelated words were not orthographically or semantically similar to the words presented in the study part of the experiment. The words that were not presented at study (homographs and unrelated words) only differed in length by 1 or 2 letters from the words presented at study. In addition, the first letter of the non-presented words was the same (e.g., *stage* and *stake*) as the words presented at study. The recognition test words were aligned on the left side of each page and there were a total of 4 pages. The recognition test words were typed on 8.5X11 printer paper and centered with 12' Times New Roman black font (see Appendix C).

## **Procedure**

The experiment was conducted in a classroom setting and participants were seated at individual desks. The LEAP-Q was given to each participant and the participants were then instructed to complete it before the experiment began. After the participants finished the questionnaire, they were given a booklet with a list of words to study. Before the participants were given the booklet, they were told to not open the booklet until the experimenter instructs them to do so. Participants were informed that the experiment consisted of two parts, the study phase and the recognition test. They were told that in the study phase, they will be presented with

a list of words in English and Spanish. They were then told that there will be a 20 second time limit to study the words on each page during the study phase of the experiment. They were instructed to spend an estimated 2 seconds on each word so that a sufficient amount of time is given to each word. They were told that the time will be kept by the experimenter with a stopwatch and after 20 seconds lapses, they will be instructed to turn the page immediately and begin studying the words on the next page. The participants were told that no questions will be answered during the experiment and to ask any questions before the experiment begins.

Participants were instructed to begin the study phase of the experiment. When the participants were finished with the study phase of the experiment, the study booklet was collected and then they were immediately given the recognition test. Participants were instructed to respond to each word by either marking the box above the “Yes”, indicating that they had recognized the word as the same word presented in the study phase of the experiment or by marking the box above the “No”, indicating that they had not recognize the word as the same word presented in the study phase of the experiment. Participants were told that there was no time limit for their responses on the recognition test. When the participants were finished with the second part of the experiment, the recognition test booklet was collected.

## RESULTS

### **False Recognition for known and unknown words similar in orthography**

One of the purposes of this study was to investigate whether false recognition based on orthographic similarities occurs for unknown words of an unfamiliar language in English native monolinguals. It was predicted that when English native monolinguals study both known words (English words) and unknown words in an unfamiliar language (Spanish words), words similar in orthography to the studied words would be falsely recognized in both languages (English and Spanish). More specifically, Spanish homographs would be more falsely recognized than Spanish unrelated words and English homographs would be more falsely recognized than English unrelated words. This prediction was tested by using a one-way within-subjects repeated measures ANOVA where the independent variable was word types (i.e., homographs, list words, and unrelated words) and the dependent variable was the proportion of words that were falsely recognized (recognition for list words) for each word type. According to Mauchly's test of sphericity, the data violated the assumption of sphericity  $\chi^2(14) = 41.48, p < .001$ . Therefore, the Greenhouse-Geisser estimates of sphericity were used ( $\epsilon = .77$ ). Analysis with the Greenhouse-Geisser correction showed significant mean recognition proportion differences for word types,  $F(3.48, 156.64) = 148.03, p = .001, \eta^2 = .77$ . As expected, Post-hoc tests, using the Least Significant Difference (LSD), showed that English homographs ( $M = .27, SD = .17$ ) were significantly more falsely recognized than English unrelated words ( $M = .14, SD = .11$ ). This evidence supports the hypothesis that false recognition occurs for known words that are similar in orthography to studied words. Most importantly, the LSD showed that false recognition also occurred for Spanish homographs. In fact, LSD showed that Spanish homographs ( $M = .54, SD =$

.19) were significantly more falsely recognized than Spanish unrelated words ( $M = .45$ ,  $SD = .25$ ). Therefore, this evidence supports the hypothesis that unknown words in an unfamiliar language similar in orthography to studied words would be falsely recognized.

### **Comparison of word types for known and unknown words**

In addition to see whether false recognition was present or not, this study further investigated whether false recognition is more likely to occur for English words (known words) or Spanish words (unknown words). According to the predictions of the orthographic distinctiveness hypothesis, orthographically distinctive words would be better recognized and, then, would be less likely to be falsely recognized. For the participants in this study, Spanish words are more orthographically distinct than English words. Thus, Spanish list words would be better recognized than English list words and Spanish homographs would be less likely to be falsely recognized than English homographs. An alternative to the orthographic distinctiveness hypothesis is the semantics distinctiveness hypothesis. According to this hypothesis, the meaning of words (i.e., semantics) can help differentiate words similar in orthographic form (Ozubko & Joordens, 2011). Participants in this study are more likely to know the meaning of English words (their native language) than Spanish words. Therefore, they would be better at recognizing English list words than Spanish list word, and would be less likely to falsely recognize English homographs than Spanish homographs. A one-way within-subjects repeated measures ANOVA was conducted to test these predictions where the independent variable was word types and the dependent variable was the proportion of words that were falsely recognized (recognition for list words) for each word type. It was found that there was a significant difference among word types  $F(3.48, 156.64) = 148.03$ ,  $p = .001$ ,  $\eta^2 = .77$ . The LSD post-hoc test pairwise comparison was

used to determine where the significant recognition proportion differences were between each of the word types. The results of each of the word types are as follows.

### ***Homographs***

To determine whether false recognition was more likely to occur for English homographs than for Spanish homographs a one-way within-subjects repeated measures ANOVA was conducted where the independent variable was word types and the dependent variable was the proportion of words that were falsely recognized for each word type. The LSD post-hoc pairwise comparison showed a significant difference in false recognition proportions between English homographs and Spanish homographs ( $p = .001$ ). There was a greater proportion of false recognition for Spanish homographs ( $M = .54$ ,  $SD = .18$ ) than for English homographs ( $M = .27$ ,  $SD = .17$ ). Thus, false recognition was more likely to occur for Spanish words similar in orthography to studied words than for English words similar in orthography to studied words.

### ***List words***

To explore whether recognition was better for English list words than for Spanish list words or vice versa a one-way within-subjects repeated measures ANOVA was conducted where the independent variable was word types and the dependent variable was the proportion of words that were falsely recognized for each word type. According to the LSD post-hoc pairwise comparison there was a significant difference in recognition proportions between English list

words and Spanish list words ( $p = .04$ ). A greater proportion of Spanish list words were recognized ( $M = .74$ ,  $SD = .17$ ) than English list words ( $M = .79$ ,  $SD = .13$ ). Therefore, recognition was better for Spanish words than for English words.

### ***Unrelated words***

A one-way within-subjects repeated measures ANOVA was conducted where the independent variable was word types and the dependent variable was the proportion of words that were falsely recognized for each word type. Results of the LSD post-hoc pairwise comparison showed a significant difference in recognition proportions for English unrelated words and Spanish unrelated words ( $p = .001$ ). There was a greater proportion of false recognition for Spanish unrelated words ( $M = .45$ ,  $SD = .24$ ) than for English unrelated words ( $M = .14$ ,  $SD = .11$ ).



## DISCUSSION

The aim of this study was to investigate whether false recognition based on orthographic similarities is present in both English, a known language, and in Spanish, an unknown language. This study further investigated whether false recognition was more likely in a known language, English, than in an unknown language, Spanish.

At recognition, English monolinguals were presented with a list of English and Spanish words; And for each word presented at recognition, they had to decide whether they had recognized the word from the study phase or had not recognized the word from the study phase. The recognition test was composed of English and Spanish words that had been presented at study (list words), words that were not presented at study but were orthographically similar to words presented at study (homographs), and words that were not presented at study and not related to words presented at study (unrelated words). Analysis of false recognition proportions showed that the false recognition proportions for English homographs was significantly greater than false recognition proportions for English unrelated words. In addition, false recognition proportions for Spanish homographs were significantly greater than false recognition proportions for Spanish unrelated words. These results support the hypothesis that false recognition based on orthographic similarity was present not only in the known language (English) but also in the unknown language (Spanish).

In previous studies, false recognition driven by orthographic similarities has been shown to exist in the native language (Lambert, Chang, & Lin, 2001; Raser, 1972). The present study provides evidence that this is also the case for unknown words of an unfamiliar language. When

participants encode words in their native language (English), in addition to having information regarding the orthography of those words, they also have information regarding the meaning of those words. Participants are encoding information in both semantic terms (meaning) and orthographic terms. However, when participants encode words of an unknown language (Spanish), they only have information regarding the orthography of those words since they do not know the meaning. Since Spanish homographs are unknown words to participants, they were unable to use the meaning to help differentiate words at recognition. Therefore, participants had to rely only on information regarding the orthographic form of Spanish homographs to assist them at recognition (i.e., participants had to base their recognition decision on orthography). Participants made mistakes at recognition (i.e., when participants identified Spanish homographs at recognition as being a word presented at study) by relying only on orthographic information thereby resulting in subsequent false recognition. This demonstrates that orthography is a significant factor in encoding information and in the recognition decision making process.

The second purpose of this study was to investigate whether false recognition was more likely to occur for English homographs than Spanish homographs and whether recognition was more likely for English list words than for Spanish list words.

Analysis of recognition proportions of list words in English and Spanish showed that correct recognition proportions for Spanish list words were significantly greater than correct recognition proportions for English list words. This finding supports the orthographic distinctiveness hypothesis. Spanish list words were better recognized because Spanish words are unique (i.e., orthographically distinct) and uncommon or unfamiliar (i.e., low frequency) to English monolinguals. Participants used both semantic and orthographic form of English list

words to facilitate their recognition decisions. Although they only had information regarding the orthographic form of Spanish list words to assist in recognition decisions, they still correctly recognized a greater proportion of Spanish list words than English list words. This provides additional evidence that orthography plays a significant role in word recognition.

The analysis of false recognition proportions for homographs in both English and Spanish showed that false recognition proportions for Spanish homographs were significantly greater than for English homographs. This finding is not consistent with the assumptions associated with the orthographic distinctiveness hypothesis. It is important to note that in studies that compare recognition in two languages and that have found support for the orthographic distinctiveness hypothesis, only recognition, not false recognition, has been analyzed (e.g., Francis & Gutierrez, 2012). In this study, the orthographic distinctiveness hypothesis was used to make predictions regarding not only recognition but also false recognition. It was assumed that since Spanish was more orthographically distinct than English, not only recognition of Spanish list words that were presented would be better (as it was actually found), but also false recognition would be less likely for Spanish homographs than English homographs (words that were not presented at study). This assumption was not supported. It seems that the process that determine whether false recognition based on orthographic similarities is more likely in one language than in the other is better explained by the semantic distinctiveness hypothesis. During the study phase, in the case of English words (but not of Spanish words) participants encoded information regarding the meaning in addition to information regarding orthographic form. At recognition, participants were able to use their knowledge of the meaning of English words to differentiate between English words that were actually presented and English words that were not presented but were

similar in orthography to presented words. This was not the case for Spanish words. That is, for English homographs, additional knowledge of the meaning assisted participants in their recognition decisions and in rejecting words that were not actually presented. For Spanish homographs, participants only had information regarding the orthographic form to assist in their recognition decisions. As a result, participants were more likely to falsely recognize Spanish than English homographs.

This study demonstrated that people not only encode the meaning of the information to be learned but also the orthographic form and that their recognition decisions are influenced by this orthographic information whether people know the meaning of words or not. The importance of orthography in word recognition cannot be ignored.

## **APPENDIX A: LANGUAGE QUESTIONNAIRE**

## LEAP-Q / Language Questionnaire

Questionnaire #		Today's Date	
Age	Date of Birth	Male	Female

1. Is English your native language?     Yes     No
2. Do you speak a second language?     Yes     No
3. If yes, which language? \_\_\_\_\_
4. Do you read a second language?     Yes     No
5. If yes, which language? \_\_\_\_\_
6. Can you understand a spoken language other than your native language?  Yes  No  
     If yes, which language? \_\_\_\_\_
7. How many years of formal education do you have? \_\_\_\_\_
8. What year college are you in? (Freshman, Senior, etc...) \_\_\_\_\_
9. Have you ever had? (Check all that apply)
- Vision problem
- Hearing impairment
- Language disability
- Learning disability
- If you checked any of the above, please explain (Including any corrections) \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_
10. Have you ever taken Spanish courses in High School?  Yes  No  
     If yes, how many years? \_\_\_\_\_
11. Have you ever taken Spanish courses in college?  Yes  No  
     If yes, how many years? \_\_\_\_\_

If you answered yes to questions #10 or #11, on a scale from 0 to 10 (10 being highest), please circle your level of proficiency in speaking, understanding and reading in Spanish:

Category	Low	Spanish Proficiency Level										High
Speaking	0	1	2	3	4	5	6	7	8	9	10	
Understand spoken language	0	1	2	3	4	5	6	7	8	9	10	
Reading	0	1	2	3	4	5	6	7	8	9	10	

12. Are you currently exposed to Spanish?  Yes  No

If you answered yes to question #12, on a scale from 0 to 10 (10 being the highest), please rate to what extent you are currently exposed to Spanish in the following contexts:

Category	Low	Spanish Proficiency Level										High
Interacting with friends	0	1	2	3	4	5	6	7	8	9	10	
Interacting with family	0	1	2	3	4	5	6	7	8	9	10	
Watching TV	0	1	2	3	4	5	6	7	8	9	10	
Listening to radio/music	0	1	2	3	4	5	6	7	8	9	10	
Reading	0	1	2	3	4	5	6	7	8	9	10	
Language lab/self-instruction	0	1	2	3	4	5	6	7	8	9	10	
Work	0	1	2	3	4	5	6	7	8	9	10	

13. On a scale from 0 to 10 (10 being the highest), please rate your attitude toward Spanish culture. (Such as, music, food, language, customs and traditions, ect...)

Attitude toward culture	0	1	2	3	4	5	6	7	8	9	10
-------------------------	---	---	---	---	---	---	---	---	---	---	----

14. On a scale from 0 to 10 (10 being the highest), please rate your interest in learning Spanish.

Interest level	0	1	2	3	4	5	6	7	8	9	10
----------------	---	---	---	---	---	---	---	---	---	---	----



## **APPENDIX B: STUDY WORD BOOKLET**

# Study

## **Orthographic Similarity and False Recognition for Unfamiliar Words**

**Experimenter:** Jeffrey Perrotte

**Purpose:** Gain a deeper understanding of recognition memory. At the end of this study you will be debriefed about the purpose of the study and what questions the study intends to answer in conclusion.

**Procedure:** There are two parts to this experiment. In the first part you will be given a list of words to study (i.e., memorize). After the first part you will immediately be given the second part of the experiment. The second part will be another list of words where you will answer either yes or no for each word. A response of “yes” indicates that you **do** recognize the word as being previously presented in the study part of the experiment and a response of “no” will indicate that you **do not** recognize the word as being previously presented in the study part of the experiment. After you have responded to all of the words in the second part of the experiment you must remain seated until the experimenter prompts you further. When the experiment is concluded and all participants have turned in their recognition test, the experimenter will debrief you on the purpose and the significance of the study and answer any questions you may have. You will also be given a Research Experience Evaluation form where you are invited (not required) to answer the questions listed and turn it in to the Psychology main front desk.

**Credit:** SONA guidelines states that you will be given ½ credit for every 30 minutes for your participation in a face-to-face study and the expected duration for this experiment is 1 hour. Therefore, you will be given 1 credit for your participation in this study.

**Instructions:** If you have any questions now is the time to ask them. Under no circumstance will questions be answered during the experiment. If you have any questions during the experiment, wait until the end of the experiment. This is to ensure that you do not disrupt the other participants during the experiment. There are no electronic devices (e.g., phone, tablet, and laptop), papers, folders, binders, or books, allowed during the experiment. There is no talking during the experiment (this includes questions for the experimenter).

## **Study Booklet**

**Procedure for first part of the experiment:** This is the first part of the experiment. This booklet contains a list of 100 words, half in English and half in Spanish. There are 10 pages in this booklet and each page consists of 10 words. You will be given 20 seconds to study the words on each page. Spend no more than 2 seconds on each word so you can study all of the words on a page. You will be instructed to turn the page and you must do so immediately and begin studying the words on the next page. When you are finished please remain seated and wait quietly for the experimenter to instruct you further.

**Important!**

**Do not turn this page until you are instructed to do so.**

**stage**

**sample**

**hojas**

**eager**

**limpio**

**saddle**

**vende**

**monte**

**awful**

**clover**

**fiel**

**store**

**umbral**

**llega**

**trozo**

**fire**

**duty**

**ciego**

**shape**

**perform**

**vencer**

**tumba**

**trigo**

**quedar**

**slave**

**sueldo**

**nacer**

**mirror**

**frenar**

**afraid**

**assign**

**county**

**reto**

**soul**

**tangle**

**huevo**

**sabio**

**gota**

**heaven**

**calvo**

**above**

**buffer**

**seco**

**grasa**

**selva**

**wisdom**

**queso**

**driver**

**alive**

**blame**



**duelo**

**reward**

**lanza**

**alma**

**trader**

**funny**

**trampa**

**loader**

**cuenta**

**asset**

**fresca**

**mito**

**cancha**

**enjoy**

**upward**

**viudo**

**ancho**

**poner**

**rayo**

**wooden**

**sewage**

**vuelta**

**profit**

**prayer**

**pista**

**trophy**

**dealer**

**salgo**

**counter**

**damage**

**powder**

**prestar**

**lumber**

**ronda**

**tejer**

**foster**

**lawyer**

**hero**

**journey**

**pena**

**logro**

**vera**

**prize**

**chapter**

**mando**

**glare**

**rompe**

**salta**

**gafas**

**defeat**

## **APPENDIX C: RECOGNITION TEST BOOKLET**

## Recognition Test

**Procedure for the second part of the experiment:** This is the second and final part of the experiment, the recognition test. You will **not** have a time limit to complete this part of the experiment. There are a total of 80 words in this recognition test; 40 of the words will be English words and 40 will be Spanish words. Some of the words in this part of the experiment are the same words that were presented to you in the first part of the experiment and some of the words were not presented to you in the first part of the experiment. Next to each word there will be a box for “Yes” and a box for “No”. You will mark the “Yes” box if you recognize that word as the same word presented to you in the first part of the experiment. You will mark the “No” box if you **DO NOT** recognize the word as the same word presented to you during the first part of the experiment. Be sure to not leave a question unanswered and to clearly mark the box with your answer.

**Instructions:** Under no circumstance will questions be answered during the experiment. If you have any questions during the experiment, wait until the end of the experiment. There are no electronic devices (e.g., phone, tablet, and laptop), papers, folders, binders, or books, allowed during the experiment. There is no talking during the experiment (this includes questions for the experimenter). When you are finished please remain quiet and an experimenter will collect your recognition booklet. When all participants have finished and handed in their recognition booklets, you will be debriefed on the purpose and significance of the study and given a Research Experience Evaluation form in which you are invited (not required) to complete and turn in to the Psychology main front desk.

alike	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
sample	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
calvo	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
glare	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
grill	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
silva	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
abide	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
profit	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
trigo	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
lawyer	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
soil	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
vuelta	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
seno	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
buffer	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
longer	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
rejo	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
vender	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
blaze	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
trozo	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
pena	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
presa	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
wisdom	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
stake	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
slate	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
pista	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
counter	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
assert	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>



manto	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
spear	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
viudo	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
hero	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
ancho	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
quemar	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
greet	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
afraid	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
queso	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
monte	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
perform	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
tejer	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
lecho	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
rodea	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
about	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
share	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
huevo	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
cancha	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
country	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
rather	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
grata	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
dummy	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
frenar	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
prestar	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
limpio	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
jugo	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
distain	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>

cuerda	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
pozo	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
foster	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
prayer	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
sueldo	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
core	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
dueto	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
llena	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
poner	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
trampa	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
chapel	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
crudo	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
vena	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
sudar	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
enjoy	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
wooden	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
piola	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
funny	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
saddle	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
defeat	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
sewage	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
fecha	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
eager	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
dealer	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>
heavy	Yes	<input type="checkbox"/>	No	<input type="checkbox"/>

**APPENDIXD: IRB APROVAL LETTER**



University of Central Florida Institutional Review Board  
Office of Research & Commercialization  
12201 Research Parkway, Suite 501  
Orlando, Florida 32826-3246  
Telephone: 407-823-2901 or 407-882-2276

## Approval of Human Research

From: **UCF Institutional Review Board #1**  
**FWA00000351,**  
**IRB00001138**

To: **Marisol Parra and Co-PI: Jeffrey Perrotte**

Date: **April 03, 2015**

Dear Researcher:

On 4/3/2015, the IRB approved the following human participant research until 04/02/2016 inclusive:

Type of Review: UCF Initial Review Submission Form  
Project Title: Orthographic Similarity and False Recognition for Unfamiliar Words  
Investigator: Marisol Parra  
IRB Number: SBE-15-11151  
Funding Agency:  
Grant Title:  
Research ID: N/A

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30 days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form **cannot** be used to extend the approval period of a study. All forms may be completed and submitted online at <https://iris.research.ucf.edu>.

If continuing review approval is not granted before the expiration date of 04/02/2016, approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a copy of the consent form(s).

All data, including signed consent forms if applicable, must be retained and secured per protocol for a minimum of five years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained and secured per protocol. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

A handwritten signature in black ink that reads "Joanne Muratori". The signature is written in a cursive style with a large, sweeping initial "J".

IRB Manager

In the conduct of this research, you are responsible to follow the requirements of the [Investigator Manual](#). On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by: Signature applied by Joanne Muratori on 04/03/2015 04:55:24 PM EDT

## REFERENCES

- Cabeza, R., & Lennartson, E. R. (2005). False memory across languages: Implicit associative responses vs. fuzzy trace views. *Memory, 13*, 1-5.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407-428.
- Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate Recall. *Journal of Experimental Psychology, 58*, 17-22.
- Francis, W. S., & Gutiérrez, M. (2012). Bilingual recognition memory: Stronger performance but weaker levels-of-processing effects in the less fluent language. *Memory & Cognition, 40*, 496-503. doi: 0.3758/s13421-011-0163-3
- Hunt, R. R., & Elliott, J. M. (1980). The role of nonsemantic information in memory: Orthographic distinctiveness effects on retention. *Journal of Experimental Psychology, 109*, 49-74.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English* (Publication No. 0-87057-105-2). Providence, RI: Brown University Press.
- Lambert, B. L., Chang, K. Y., & Lin, S. J. (2001). Effect of orthographic and phonological Similarity on false recognition of drug names. *Social Science and Medicine, 52*, 1843-1857.
- Marian, V., Blumenfeld, H.K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*, 940-967.

Ozubko, J. D., & Joordens, S. (2011). The similarities (and familiarities) of pseudowords and extremely high-frequency words: Examining a familiarity-based explanation of the pseudoword effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 123-139.

Parra, M. (2013). *False recognition driven by meaning and form: The dynamics of bilingual memory representations*. (Doctoral dissertation) Florida Atlantic University.

Raser, G. (1972). False recognition as a function of encoding dimension and lag. *Journal of*

*Experimental Psychology*, *93*, 333-337. Roediger, H. L. III., & McDermott, K. B. (1995).

Creating false memories: Remembering words not presented in lists. *Journal of*

*Experimental Psychology: Learning, Memory and*

*Cognition*, *21*, 803-814.

SPANISH ROYAL ACADEMY: Database (CREA) [online] *Reference Corpus Current Spanish*

Retrieved from <http://www.rae.es/creanet.html>