# I KNOW SPANISH. WHICH ROMANCE LANGUAGE SHOULD I LEARN NEXT?: L2 INFLUENCE ON L3 WORD RECOGNITION

David Beard, Emily Neff, and Ally Bucher

Miami University
bearddj@miamioh.edu

## ABSTRACT

The present study addresses historical and contemporary discussions on second and third language vocabulary development, specifically the effects of lexical similarity. Second language learners of Spanish aspiring to learn another romance language completed a translation recognition task with Spanish as the source language and Portuguese or Italian as the target language. Cognate and non-cognate targets were used. Reaction times were modeled as a function of language and word type using a linear mixed-effects model with subject and item serving as random effects. Results show that there is no significant processing advantage in recognizing words in Italian or Portuguese. With respect to word type, cognates were responded to significantly faster than non-cognates. These results speak to historical and contemporary accounts of lexical similarity in that Spanish and Italian are rated at 82% and Spanish and Portuguese at 89% [13]. Implications speak to both language researchers and educators.

**Keywords**: second language acquisition, third language acquisition, vocabulary, cognate facilitation effect, lexical similarity measures

**Acknowledgements:** The authors would like to extend our sincere gratitude to the Department of Spanish and Portuguese at Miami University for their financial support as well as the Statistics Consulting Center at Miami University for their expertise and consultation.

## 1. INTRODUCTION

Everyone wants to know more languages but learning languages as an adult is no small feat [5]. It takes years of study to reach high levels of proficiency [5]. This combination of desire to learn and knowledge that it takes a long time to do so leads our brains to wonder how best to go about this process. We are wise to look for familiar patterns in our new language environment because we know that similarities in lexical structure between source and target languages can facilitate target language development [5,12]. To that end, the present study addresses the effects these similarities present to someone who knows Spanish and aspires to learn either Italian or Portuguese.

## 2. MOTIVATION / BACKGROUND

### 2.1 Language learners

Faculty members in World Languages departments want effective curricula that encourage developing language proficiency in their students [4]. Graduate students in these departments often need to take translation exams to demonstrate their ability to read and understand material relevant to their research in other languages. Many highly motivated undergraduate students majoring or minoring in these departments are completing their degree in a second language and some begin a third prior to graduation. They know that employers are looking for multilingual candidates [5]. They are busy and want to be efficient learners, asking questions that prompt titles of articles like this one, such as, "If I know Spanish, which romance language should I learn next?" Furthermore, it is not uncommon to hear popular references about romance languages and their intimate relation, like Homer Simpson's famous "Italian and Spanish are the same language" [14] and Dee, from It's Always Sunny in Philadelphia, claiming that "if you know Latin, you know like five languages" [11] as she is about to read instructions in Spanish.

### 2.2 Historical and Psycholinguistic Approaches

What motivates the aforementioned discussion, and the choice of languages people use to compare, is the fact that many words in the compared languages look and/or sound and mean alike as well, (e.g. English: paper, Spanish: *papel*). These similarities have not only affected popular discussion on the topic but have also been utilized in studies by historical linguists and psycholinguists [8,9].

Historical linguists have constructed lexical similarity charts to show the percentage of vocabulary two

languages share. For example, Spanish and Italian share 82% of their lexicons while Spanish and Portuguese share 89% [13]. This measure comes from historical documents and uses etymology to assess the similarity of lexicons, essentially providing insight into their shared orthographic and semantic history [8,13].

Psycholinguists have conducted experiments and have accumulated a growing list of significant predictor variables describing what it takes to recognize and produce a word, (e.g. frequency, word length, concreteness, imageability, etc.) [9]. Within experiments that specifically address bilinguals, much work has been done that has looked at the effects of word similarities between languages [6,9,12], but few psycholinguistic studies have engaged historical accounts.

### 2.3 The cognate facilitation effect (CFE)

A cognate is the term for these translations that look and/or sound similar. It is defined here as a word translation that shares orthography and/or phonology with its counterpart, (e.g. *pared – parede – parete* in Spanish, Portuguese and Italian respectively). The cognate facilitation effect (CFE) states that cognates will be processed faster than non-cognates and came about after decades of empirical research [9,12]. Current debates about the nature of this facilitatory effect hinge on whether the effect is caused by increased form overlap found in cognates, (i.e. shared orthography and/or phonology) or whether this effect is caused by increased structural overlap (i.e. shared morphology) [12]. This is a well attested effect and it has been replicated in many studies [6,9,12]. World language teachers have also been trained for years to utilize cognates in the classroom to exploit productive use of first language resources and there are several cognate based vocabulary activities available in the L2 pedagogy literature [5].

Finally, it is worth noting that languages share cognates because of language contact and phylogenetic history. Beyond orthographic similarity, we also observe likeness in other levels of linguistic analysis, (e.g. phonemic inventory, phonotactic rules, word formation rules, syntactic and semantic categories and pragmatics) and it is these similarities that facilitate lexical development beyond the set of cognates that two languages share.

### 2.4 Research Questions

This talk of effective language learning and our knowledge of the CFE lead us to address two questions:

1. To an L2 learner of Spanish who wants to learn more romance languages, is it more effortful to process Italian or Portuguese orthography?
2. Is there a Cognate Facilitation Effect (CFE) present in using Spanish to recognize these other two languages?

## 3. METHODOLOGY

### 3.1 Participants

62 participants (mean age = 20 years old, SD = 1.1) were recruited from the undergraduate student body at Miami University, USA. They were randomly assigned to a language (i.e. Italian or Portuguese) with 31 participants in each group. Participants were paid and provided informed consent in accordance with the institutional review board requirements. All participants had normal or corrected to-normal vision and no self-reported language-related disorders. 61/62 participants identified as White, 1/62 as African-American, 55/62 as women and 7/62 as men. All participants grew up in the United States, were first language learners of English and second language learners of Spanish, and all had aspirations to learn another romance language. Per the ACTFL (American Council on the Teaching of Foreign Languages) proficiency guidelines [1], participants had an average second language proficiency level of intermediate-mid. All participants were Spanish majors or minors and were currently enrolled in an upper-division university level course in Spanish. Through the use of a pre-screening questionnaire, participants were excluded if they spoke Spanish at home, were not currently enrolled in a Spanish class, or if they had begun learning any other romance language.

### 3.2 Stimuli

60 words in Spanish were selected and would be used as primes in the translation recognition task. 30 of these words are cognates with Italian and Portuguese and 30 of them are not cognates with either language. To minimize interference from participants' L1, none of the 60 words in Spanish is a cognate with English. Word lists were balanced for frequency (cognate mean frequency per million = 20.7, SD = 15.2, non-cognate mean frequency per million = 14.2, SD = 14.3, t-test of cognate and non-cognate frequency = 0.09, $p > 0.05$), and word length (cognate mean word length = 6.2 letters, SD = 2.1, non-cognate mean word length = 6.5 letters, SD = 1.5, t-test of cognate and non-cognate word

length = 0.49, p = > 0.05). Lexical measures were obtained using EsPal [7]. Additionally, 60 non-words were generated using Wuggy [10]. They would appear on the target screen with Italian or Portuguese words and thus were matched in length and syllable structure to the appropriate target. To ensure that participants could not rely on form overlap between the prime and target to predict their response, half of the non-words appeared to be cognates with the Spanish prime and they were paired with their appropriate Italian or Portuguese non-cognate targets.

### 3.3 Design

A translation recognition task was utilized to address our research questions. This task was chosen because it is a valid measure to address lexical access [9] and because it approximates (classroom) word learning when translations are used. A translation recognition task is similar to a double lexical decision task that utilizes priming. Participants are seated at a computer and use a button box to respond. They are presented with a fixation cross in the center of the screen followed by the prime in Spanish for 500 milliseconds. Immediately after, the target screen appears, and it contains two strings of letters on the screen, one aligned left and one aligned right. One is the correct translation of the prime and the other is a well-formed non-word in the appropriate target language. Participants respond as quickly and accurately as possible using a button box to select the string of letters they believe to be the correct translation of the prime. Reaction time and accuracy are recorded. All computer-based experiments were programmed using E-Prime 3.0. The logic is that if the prime is more related to the target, it will facilitate a faster reaction time [9].

### 3.4 Procedure

Half of the participants were given a list of words in Italian and the other half in Portuguese, both with their translations in Spanish. Participants were instructed to study the list for a period of five minutes. Participants then completed the translation recognition task in the appropriate language, (i.e. Italian or Portuguese). This task contained the same words participants saw at the beginning of the session.

## 4. RESULTS

Incorrect responses were eliminated and then reaction times (RTs) were analyzed using a linear mixed-effects model with a random slope for "word type" by subject and a random slope for "language" by item. This was accomplished using the statistics software R, the package 'languageR', [2] and the 'lmer' function [3]. Effect and null models were constructed to test for main effects and interactions of language and word type, and likelihood ratio tests were used to compare them. We found no significant effect for language ($\chi^2$ (1) = 0.38, p > 0.05) but we did find a significant effect for word type ($\chi^2$ (1) = 48.27, p < 0.05). In adding in the interaction between language and word type we found no significant effect ($\chi^2$ (1) = 0.80, p > 0.05). Mean accuracy by language was 0.83 and 0.79 for Italian and Portuguese respectively and mean accuracy by word type was 0.93 and 0.69 for cognates and non-cognates respectively.

Figures 1 and 2 show mean differences in raw RTs by language and word type. As we can see, mean RT difference between Italian and Portuguese is only 17ms, with RTs to Portuguese being slightly faster. However, mean RT between cognates and non-cognates is 181.5ms.

**Figure 1**: Mean differences in raw reaction time (RT) by language, ($\chi^2$ (1) = 0.38, p > 0.05).

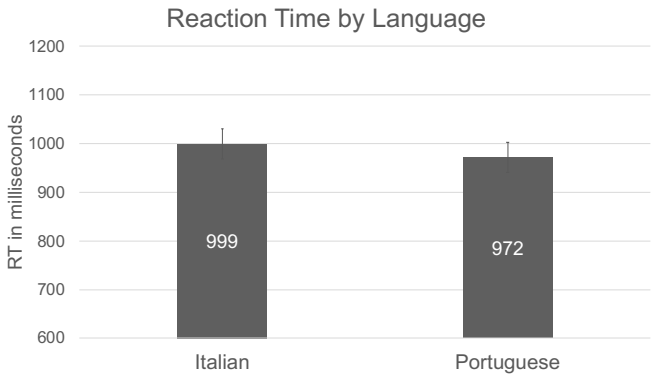
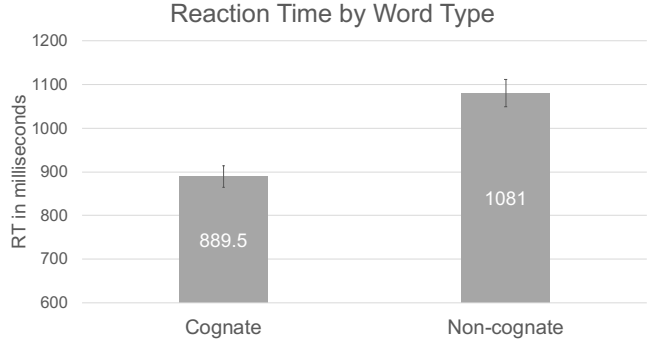



**Figure 2**: Mean differences in raw reaction time (RT) by word type, ($\chi^2$ (1) = 48.27, p < 0.05).



## 5. DISCUSSION AND CONCLUSIONS

To address our first research question, our results suggest that there is not a significant difference in processing Italian and Portuguese orthographies if one knows Spanish. However, we cannot conclude that this null effect demonstrates that there is no difference in reality. Perhaps it is the case that both languages are too similar to Spanish to note a significant difference, but we estimate that we would find these types of differences in languages with greater divergence in the lexicons with the same experimental design, (e.g. Spanish to Hungarian).

To address our second question, we can answer with confidence that there is. This is not surprising as the CFE has been replicated in many studies [9,12]. The interesting factor here is that this is at a very early stage of L3 acquisition and that the facilitation is coming from the learners' L2. The increased orthographic overlap found in cognates as opposed to non-cognates probably facilitates faster reactions times because the orthographic and semantic representations in the Spanish prime are still active upon observing the target and when the target is a cognate, the orthographic overlap in the target connects with the prime and speeds one's reaction. On the other hand, this is not the case with non-cognates as they lack the orthographic connection needed to benefit from the prime.

What can we take away from this? Our results indicate that L2 learners of Spanish do not process the orthography of Italian or Portuguese distinctly. We witnessed only observably faster RTs in Portuguese, which does align with historical accounts of lexical similarity in that Spanish and Italian are rated at 82% while Spanish and Portuguese are rated at 89% [13].

We also found that cognates do offer a significant advantage in processing over their non-cognate counterparts irrespective of the language in which the learner finds them. This indicates that languages with similar lexicons may offer some advantages in word recognition to the L2 learner, but this still needs to be tested.

In closing, we would like to point out that while lexical similarity and cognates may or may not be significant factors in learning an L2 or L3 in the long run, in order for aspiring learners to reach their ultimate language learning goals, it is our opinion that they focus their attention on other potentially more significant variables like their motivation to learn a particular language, their heritage, their career, or meeting the love of their life.

## REFERENCES


1. American Council on the Teaching of Foreign Languages (2012). *Proficiency Guidelines.* Retrieved from: https://www.actfl.org/sites/default/files/pdfs/public/ACTFLProficiencyGuidelines2012_FINAL.pdf
2. Baayen, R. H. (2013). LanguageR: Data sets and functions with *Analyzing linguistic data: A practical introduction to statistics.* R package version 1.4.1. Available from https://CRAN.R-project.org/package=languageR
3. Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1-48.
4. Beard, D. & Castañeda, M. (2019). Beyond Input: Promoting Interpersonal Communication with Task-Based Language Teaching. *CSCTFL Report* (In Press).
5. Blake, R. J. & Zyzik, E. C. (2016). *El español y la lingüística aplicada.* Georgetown University Press: Washington, DC.
6. Dijkstra, A., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. Journal of Memory and Language, 41, 496-518.
7. Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop Shopping for Spanish Word Properties. *Behavior Research Methods*, 45, 1246-1258.
8. Goddard, I. (1975). Algonquian, Wiyot, and Yurok: Proving a Distant Genetic Relationship. In M. D. Kinkade, K. L. Hale, & O. Werner (Eds.), *Linguistics and anthropology in honor oJe. F. Voegelin* (pp. 249-262). Lisse: Peter de Ridder Press.
9. Grosjean, F. (2013). *The psycholinguistics of bilingualism.* Oxford, UK: Blackwell Publishing.
10. Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. Behavior Research Methods 42(3), 627-633.
11. Rosell, R. (Writer), Marder, S. (Writer), & Shakman, M. (Director) (October 30th, 2008). Sweet Dee has a heart attack [Television series episode]. In C. Day, *It's Always Sunny in Philadelphia.* Philadelphia, PA: RCG Productions.
12. Sánchez-Casas, R. & García-Albea, J. E. (2005). The Representation of Cognate and Noncognate Words in Bilingual Memory: Can Cognate Status Be Characterized as a Special Kind of Morphological Relation? In J. F. Kroll and A. M. B. De Groot (eds.), Handbook of bilingualism: Psycholinguistic approaches, pp. 226–250. New York: Oxford University Press.
13. Simons, G. F., & Fennig, C. D. (eds.) (2018) *Ethnologue: Languages of the world,* 21st ed. Dallas, Texas: SIL International. Online version available from http://www.ethnologue.com
14. Swartzwelder, J. (Writer), & Kirkland, M. (Director). (December 3rd, 2000). The computer wore menace shoes [Television series episode]. In M. Groening (Producer), *The Simpsons*. 20th Century Fox Television.