2008

Using cognates to investigate cross-language competition in second language processing

Gretchen Sunderman
Ana I Schwartz, *University of Texas at El Paso*

Using Cognates to Investigate Cross-Language Competition in Second Language Processing

GRETCHEN SUNDERMAN
Florida State University
Tallahassee, Florida, United States

ANA I. SCHWARTZ
University of Texas at El Paso
El Paso, Texas, United States

■ In the psycholinguistic literature, cognates are often defined as words that share aspects of spelling, sound, and meaning across languages (e.g. *piano* in Spanish and *piano* in English).[1] Second language (L2) learners are often quick to take advantage of the similarities found in cognates and transfer knowledge from their first language (L1) to facilitate vocabulary acquisition and understanding in the L2. Unfortunately, words

[1] In linguistics, *cognates* are often defined as words that share a common etymological origin. Here, we employ instead a standard definition of cognates used within the psycholinguistic literature, namely any two words with shared aspects of spelling, sound, and meaning across two languages. Within psycholinguistic research, the focus is on language processing. Only aspects of the underlying mental representation (orthography, phonology, and semantics) are hypothesized to affect processing. Historical linguistic roots of a word would not be represented in the mind of the language user, and thus the linguistic definition of cognates based on etymological roots does not capture the form–meaning overlap that psycholinguists are interested in. See Carroll (1992) for a similar argument.

that are seemingly similar across languages may in fact be quite different. Sometimes learners encounter words in an L2 that are similar in form to those in their L1, but do not share meaning (i.e. false friends such as *embarazada* meaning "pregnant" in Spanish and not "embarrassed"). At other times, learners encounter words that are similar in form but differ in the degree to which they share meaning. Take the orthographically identical cognate pair *grave/grave* in English and Spanish. In both English and Spanish, *grave* means "serious." However, in English alone *grave* also means "burial place." Therefore, what distinguishes *partial cognates* of this kind from full cognates is the existence of another, unshared, homographic meaning in L2.

In this article, we present a psycholinguistic study investigating how the semantic ambiguity found in partial cognates affects lexical processing. The assumptions we adopt are those that underpin a well-known psycholinguistic model of the bilingual lexicon, the bilingual interactive activation (BIA) model (Dijkstra & Van Heuven, 1998; Dijkstra, Van Jaarsveld, & Ten Brinke, 1998; Van Heuven, Dijkstra, & Grainger, 1998). The BIA model represents visual word recognition by bilinguals. It is based on the notion of *competition*: The sight of a string of letters on the page excites a number of words in the reader's mind which compete as to which provides the best match for what has been seen. What is striking about the BIA model is the assumption that words from both the user's languages participate in this competition process, at the feature, letter, and word levels. The model posits a single formal store, which holds word forms for both languages. It is only at a relatively advanced stage of the competition process that language-specific criteria are brought to bear. Language-level nodes specify language membership (in the original model, Dutch versus English) and have the ability to suppress all of the activated words belonging to the other language. The assumption of the BIA model is thus that, when a proficient bilingual sees a letter string, several words may be activated, regardless of language. For example, according to the model, when a Spanish–English bilingual reads the letter string *gate,* the letter string *gato* (cat) is also activated and competing for selection. Assume that the language user in question is reading in English. Once a match between the letter string *gate* and the word *gate* is established, the word sends activation to the English language node, which then shuts off the candidates from the other language.

An extensive body of research supports the predictions of the BIA model and shows that bilinguals activate information about words in both languages in parallel, regardless of their intention to function within one language alone (See Dijkstra 2005 for a review of bilingual word recognition). Some of the evidence for the BIA model comes from studies using cognates. This work (e.g., Dijkstra et al., 1998; Van Hell & Dijkstra, 2002) has consistently demonstrated that bilinguals are faster at

recognizing words that are cognates than at recognizing noncognates, an effect known as *facilitation*. It is not surprising that monolinguals show none of these effects. In other words, when bilinguals access their mental lexicons to retrieve a word that is a cognate, the overlap of form and meaning between the two languages speeds the retrieval process. Thus, shared meaning or convergence in meaning across languages helps performance.

The opposite pattern has been found for words whose meanings across languages diverge. They include words that have multiple translations, such as the polysemous English word *glass*, which translates into two words in Spanish: *vidrio*, as in the material of glass, and *vaso*, as in something to drink out of (Elston-Güttler, 2000; Tokowicz 2001; Tokowicz & Kroll, 2007). Tokowicz found that bilinguals were slower to translate words such as this with multiple meanings, suggesting that the two meanings were active and competing for selection and that this slowed translation.

We therefore see that although the cognate's cross-language overlap in lexical form can facilitate bilingual word recognition, any divergence or nonequivalence in meaning can slow processing. By manipulating cognate status and meaning overlap in the current study, we were able to separately examine the effects of cognate facilitation and semantic competition. We predicted that the cognate facilitation observed so frequently and consistently in previous research would be attenuated or possibly eliminated for partial cognates because of the competition from multiple meaning mappings.

METHOD

To test this prediction, late Spanish–English bilinguals (i.e., adult learners of English) completed a visual lexical decision task ("decide whether this letter string is a word or not") which required the identification of two types of cognate: (a) cognates for which there is no semantic ambiguity; they map on to just one meaning across the two languages (e.g., *piano*) or (b) cognates with semantic ambiguity for which there is more than one meaning (e.g., *grave*). We refer to these two types of cognates as *full* and *partial*, respectively. It is important to note that the term *partial cognate* here refers only to the overlap in meaning, not to an overlap in form.

The logic in using the lexical decision (LD) task is that it requires individuals to search their lexicons for a lexical representation that matches the letter string presented. The representation of a lexical item contains information regarding the word's orthography, phonology, and

semantics, and these aspects of the word are retrieved during the task. Although it is possible to recognize a visually presented word based solely on its orthographic information, there is much evidence demonstrating that semantic information is also activated (see Neely, 1991, for a review). Thus, the LD task is an effective tool for examining how words compete with each other at the semantic level.

Participants

Twenty-one participants from the University of Valencia, Spain, participated in the experiment. All participants were native speakers of Spanish who were proficient in English. The participants' ages ranged from 18 to 31, with a mean age of 24.5 years. The participants completed a language history questionnaire in which they were asked to self-assess their proficiency in reading, writing, speaking, and listening in English and Spanish on a scale of 1 to 10. They also reported any immersion experiences they had in English-speaking countries as well as their familiarity and proficiency with languages other than English and Spanish. On average the participants had been studying English for 12.1 years, starting at around the age of 10. They had spent an average of 4.1 months in an English-speaking country. Of the 21 participants, 13 reported having high proficiency in Valenciano as well as in Spanish and English. Of these 13, 4 considered themselves native speakers of Valenciano. Thus, the majority of the participants were trilingual, with proficiency in Spanish, Valenciano, and English. Their self-assessed ratings indicated that although they considered themselves relatively proficient in English (mean rating: 6.3) they were dominant in Spanish (mean rating: 9.6). All participants were paid for their participation in the study.

Materials

The target words consisted of 42 full cognates[2] (e.g., *piano*) and 38 partial cognates (e.g., *grave*). For each critical word, a noncognate control word matched in terms of word frequency and length was selected. An additional 48 noncognate fillers were included. Also added was a set

[2] The full cognate items were selected from a previous study (Schwartz, Kroll, & Diaz, 2007). In this study, bilinguals assessed the phonological similarity of cognate pairs. Because the objective of this article is to examine the influence of cognate status and cross-language semantic overlap, we employ those cognates that were rated as being highly phonologically similar (n = 22) to reduce any potential confounds due to limited phonological overlap across languages.

of nonword distractors. The nonwords were pronounceable letter strings created by changing one or two letters from an actual English word (e.g., *garrons*). The materials were divided into two lists (A and B). See Table 1 for an illustration of the materials used. We stress again that the classification of *partial cognate* is based on semantic ambiguity and does not relate to formal aspects of the word. Both the full and partial cognates had varying degrees of orthographic overlap.

Procedure

Participants were recruited through announcements posted across the university campus. When they arrived at the laboratory, they were greeted in English (L2). After completing an informed consent form (written in Spanish to ensure that the experimental description was clear), participants were seated in front of a laptop computer. Instructions were presented on the laptop display in English. These instructions were also read aloud to the participants. Clarifications were given in Spanish when needed. Participants were instructed that they would see letter strings presented, one at a time, in the middle of the computer screen. They were told to decide as quickly and accurately as possible whether each letter string formed an actual English word or not and to signal their answer by pressing the appropriate key on the response box. If they were unsure, they were told to guess. Participants completed 10 practice trials before the experimental phase began, after which the experimenter left the room.

Each trial was initiated by the presentation of a fixation point (+) in the center of the screen. It remained on the screen until the participant pressed the center key on the response box. A word or nonword was then presented and disappeared after the participant had made a key response or 3 seconds had elapsed. Participants pressed the left-hand key

TABLE 1
Examples of Stimuli and Their Lexical Characteristics

	Critical words			Control words		
	Examples	Frequency*	Length	Examples	Frequency*	Length
Full cognates	*band* *actor* *calm*	41.9	6.2	*bond* *affair* *coat*	40.0	6.3
Partial cognates	*grave* *rest* *labor*	77.8	5.5	*gift* *road* *Letter*	75.3	5.6

* Frequency in English from Kucera and Francis (1967).

on the response box to indicate "no" for nonwords and the right-hand key to indicate "yes" for actual words.

RESULTS

The strength of the LD task, as widely used in psycholinguistic research, is that it provides two measures of a participant's ability to recognize a given word. The first relates to the accuracy with which the word has been recognized. The second (more informative in the present case) relates to *reaction time* (RT): how long the participant takes to decide that a string of letters is, in fact, an actual word. Longer reaction times and/or decreases in accuracy are assumed to reflect heavier cognitive demands on the reader. In the context of the current study, it was assumed that both accuracy and reaction time indicated how relatively difficult it was for participants to recognize words that were full cognates, partial cognates, or noncognate controls. The hypothesis was that RT would be slower and accuracy less when the target words were partial cognates.

We report reaction time and accuracy separately. Mean RTs for correct trials were calculated for each participant in each condition. RTs that were faster than 300 milliseconds (ms) or slower than 3,000 ms were counted as outliers and excluded from the analyses. RTs that were more than 2.5 standard deviations above or below the participants mean RT were also considered outliers and eliminated from analyses. The standard deviation cut-offs were calculated separately for the words and nonwords. These data trimming procedures led to an exclusion of 2% of all trials.

Reaction Time Data

Mean lexical decision RTs and mean percent accuracy for each of the four conditions are reported in Table 2. Overall, participants were faster at recognizing cognates (669.3 ms) relative to noncognates (706.9 ms). This result is consistent with prior research and suggests that the activa-

TABLE 2
Reaction Time (Milliseconds) and Accuracy (%) Means for Full and Partial Cognates

	Full cognates		Partial cognates	
	RT (ms)	Accuracy (%)	*RT* (ms)	Accuracy (%)
Cognates	671	90	667	92
Noncognate controls	738	73	676	86
Magnitude of facilitation	67	17	9	6

tion of the cognates' representations across languages allowed for quicker recognition of the cognate words. To determine whether this cognate facilitation was significantly reliable, the mean RTs were submitted to a two-way repeated measures ANOVA. The variables were cognate status (cognate vs. noncognate) × meaning overlap (full vs. partial).[3] The analysis revealed a main effect of cognate status, $F(1, 20) = 19.18$, $MSE = 1541.67$, $p < .001$, which verified the statistical reliability of the cognate benefit.

Table 2 also shows that the overall magnitude of cognate facilitation was greater for the full cognates relative to the partial cognates. More specifically, full cognates were recognized on average 67 ms faster than their noncognate matched controls, whereas partial cognates were recognized only 9 ms faster than their respective controls. Indeed, the main effect of cognate status from the ANOVA just mentioned was qualified by a significant two-way interaction with meaning overlap, $F(1, 20) = 6.85$, $MSE = 2566.5$, $p < .05$, thus suggesting that the semantic ambiguity of partial cognates was affecting processing.

Accuracy Data

The pattern observed in the accuracy data closely paralleled that of the latency data described earlier. Cognates were more accurately identified (91%) than noncognates (80%), so once again, cognate status facilitated performance. To determine whether this cognate facilitation was significantly reliable, the mean accuracy rates were submitted to a second two-way repeated measures ANOVA. Once again, the variables were cognate status (cognate vs. noncognate) × meaning overlap (full vs. partial). The analysis revealed a main effect of cognate status, $F(1, 20) = 17.16$, $MSE = 169.18$, $p < .01$, which verified the statistical reliability of the cognate facilitation. This result provides further evidence that the coactivation of the cognates' representations across languages facilitated recognition.

The RTs reported in Table 2 also reveal that, as with the RT data, the overall magnitude of cognate facilitation was greater for the full cognates relative to the partial cognates. More specifically, full cognates were recognized on average 17% more accurately than their matched controls, whereas partial cognates were recognized only 6% more accurately than their respective controls. Indeed, the main effect of cognate status

[3] It is important to note that *meaning overlap* applied only to the cognates and not the noncognate controls. For the purpose of the analyses, the noncognate controls were simply coded as either *full* or *partial* to correspond with their respective full or partial critical cognate pair in order to be matched in terms of frequency and length.

from the ANOVA cited earlier was qualified by a significant two-way interaction with meaning overlap, $F(1, 20) = 9.64$, $MSE = 61.05$, $p < .05$. Again, this suggests that the additional meaning mapping of the partial cognates is affecting accuracy.

DISCUSSION

Consistent with previous findings in the literature, and as expected, we found that individuals processed cognates more quickly and accurately than they processed noncognate control words. The overlap of form and meaning across languages for the cognates facilitated lexical access. This cognate effect was also found for partial cognates in which the meaning overlap was not perfect. Thus, even though the meaning overlap was not as strong in the partial cognates as in the full cognates, the partial cognates were nevertheless retrieved more quickly and accurately than noncognate control words. Therefore, this result extends the cognate effect to include partial cognates.

The interaction between cognate status and meaning overlap in both the RT and accuracy data suggests that partial cognates are indeed processed differently than full cognates. The incomplete semantic overlap of the partial cognates caused these items to be processed more slowly and less accurately compared with the full cognates. The result indicates that imperfect meaning overlap at the semantic level can reduce the general facilitative effects of lexical form; it is consistent with prior work demonstrating that cognate facilitation can be eliminated or reversed when there is not complete lexical overlap (Dijkstra, Grainger, & van Heuven, 1999; Schwartz, Kroll, & Diaz, 2007; Tokowicz & Kroll, 2007). Overall, the results indicate some processing costs associated with the semantic ambiguity found in partial cognates.

PEDAGOGICAL IMPLICATIONS

An important difference between proficient and nonproficient bilingual readers lies in how readers determine the meaning of unknown words (Jimenez, García, & Pearson, 1994). One strategy for improving reading comprehension is structured practice recognizing and utilizing cognates across languages (Moss, 1992). Indeed, research with Spanish–English bilingual biliterate children found a strong relationship between their L2 English reading comprehension and their ability to recognize cognates (Nagy, García, Durngunuğlu, & Hancin-Bhatt, 1993). Although Nagy et al. suggest that cognate instruction might facilitate L2 reading,

they speculate that a learner's ability to deal with cognates with partial semantic overlap may require a "higher level of cognitive flexibility and metacognitive awareness" (p. 254). Given the results of the current study, namely that partial cognates do impose an increased processing load compared with full cognates, we recommend that teachers use explicit instruction to draw learners' attention to this particular type of cognate relationship.

We also suggest that when teaching cognate forms for reading, L2 teachers focus not only on building strong connections to meanings that are shared across L1 and L2 but also to meanings that are not shared. Vocabulary teachers might first focus on shared semantic values before introducing values that are not shared. In the case of values that are not shared, learners will be faced with the competing meaning representations of the cognate and will be forced to reconcile themselves gradually to the need to make a choice, which may be contextually determined. This approach will add depth to the lexical representation and enable learners to identify the differing contexts that support the shared and nonshared meanings of the partial cognate. Recent work (Perfetti & Hart, 2001) suggests that how easy it is to identify a given word depends not only on the word itself, but also on factors related to the reader and the context. In other words, as learners become exposed to the varying contexts in which the two competing meanings are used, they will be able to build links to both meanings, which can be quickly identified and retrieved from the lexicon. They will thus become stronger L2 readers.

ACKNOWLEDGMENT

We express our gratitude to John Field for serving as guest editor of this special issue. We also thank Michael Leeser, Christopher Reenock, and the anonymous reviewers for their helpful comments. This research was supported by a grant from the National Institutes of Mental Health (Grant 1 F31 MH66476-01A1) and the National Science Foundation (Grant BCS 0212571).

AUTHORS

Gretchen Sunderman is an assistant professor of Spanish in the Department of Modern Languages and Linguistics at Florida State University, Tallahassee, Florida, United States. Her research interests include second language lexical processing and the role of individual differences in second language acquisition.

Ana I. Schwartz is an assistant professor of psychology at the University of Texas at El Paso, United States. Her research interests include bilingualism, the cognitive architecture of the lexicon, and individual differences in reading fluency and skill.

REFERENCES

Carroll, S. (1992). On cognates. *Second Language Research, 8,* 93–119.

Dijkstra, T., 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.

Dijkstra, T., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition, 1,* 51–66.

Dijkstra, A., & Van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 189–225). Hillsdale, NJ: Erlbaum.

Dijkstra, A. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). New York: Oxford University Press.

Elston-Güttler, K. (2000). *An enquiry into cross-language differences in lexical-conceptual relationships and their effect on L2 lexical processing.* Unpublished doctoral dissertation, University of Cambridge, England.

Jimenez, R. T., García, G. E., & Pearson, P. D. (1994). Three children, two languages and strategic reading: Case studies in bilingual/monolingual reading. *American Educational Research Journal, 32,* 67–98.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Moss, G. (1992). Cognate recognition: Its implications in the teaching of ESP reading courses to Spanish speakers. *English for Specific Purposes, 11,* 141–158.

Nagy, W., García, G., Durngunuğlu, A., & Hancin-Bhatt, B. (1993). Spanish–English bilingual students' use of cognates in English reading. *Journal of Reading Behavior, 25,* 241–259.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.

Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skill. In D. S. Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity. Decade of behavior* (pp. 67–86). Washington, DC: American Psychological Association.

Schwartz, A. I., Kroll, J. F., & Diaz, M. (2007). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Language and Cognitive Processes, 22,* 106–129.

Tokowicz, N. (2001). *Meaning representation within and across languages.* Unpublished doctoral dissertation, Pennsylvania State University, University Park, PA, United States.

Tokowicz, N., & Kroll, J. F. (2007). Number of meanings and concreteness: Consequences of ambiguity within and across languages. *Language and Cognitive Processes, 22,* 727–779.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9,* 780–789.

Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39,* 458–483.