

Translation norms for English and Spanish: The role of lexical variables, word class, and L2 proficiency in negotiating translation ambiguity

ANAT PRIOR AND BRIAN MACWHINNEY
Carnegie Mellon University, Pittsburgh, Pennsylvania

AND

JUDITH F. KROLL
Pennsylvania State University, University Park, Pennsylvania

We present a set of translation norms for 670 English and 760 Spanish nouns, verbs and class ambiguous items that varied in their lexical properties in both languages, collected from 80 bilingual participants. Half of the words in each language received more than a single translation across participants. Cue word frequency and imageability were both negatively correlated with number of translations. Word class predicted number of translations: Nouns had fewer translations than did verbs, which had fewer translations than class-ambiguous items. The translation probability of specific responses was positively correlated with target word frequency and imageability, and with its form overlap with the cue word. Translation choice was modulated by L2 proficiency: Less proficient bilinguals tended to produce lower probability translations than more proficient bilinguals, but only in forward translation, from L1 to L2. These findings highlight the importance of translation ambiguity as a factor influencing bilingual representation and performance. The norms can also provide an important resource to assist researchers in the selection of experimental materials for studies of bilingual and monolingual language performance. These norms may be downloaded from www.psychonomic.org/archive.

The study of bilingual language processing exploits the presence of translation equivalents in the bilingual speaker's two languages as an important research tool. The impact of different lexical characteristics in both the source and the target language on performance in translation recognition and production has been taken as evidence regarding the organization of the bilingual lexicon and conceptual system (e.g., De Groot, 1992; Kroll & Tokowicz, 2005; Sánchez-Casas & García-Albea, 2005). The translation task itself allows one to study the dynamic online interaction between the two language systems and to examine specific cross-linguistic relations between words.

Furthermore, studies using the picture-word Stroop task consider the way in which the presence of a translation in one language affects production in the other language (e.g., Costa, Miozzo, & Caramazza, 1999; Hermans, Bongaerts, De Bot, & Schreuder, 1998). In studies of bilingual word recognition, cross-language priming across translation equivalents has also been used to determine the degree to which aspects of lexical form and meaning are shared for a bilingual's two languages even when the two languages do not share the same script (e.g., Chen & Ng, 1989; Gollan, Forster, & Frost, 1997; Jiang, 2000; Keatley, Spinks, & De Gelder, 1994). In addition, researchers interested in general mechanisms of language

production have started to use translation tasks as a tool that allows them to avoid some of the shortcomings of picture naming, which is necessarily restricted to a limited set of concrete nouns (e.g., Gummior, Bolte, & Zwitserlood, 2006; Jescheniak & Levelt, 1994; Vigliocco, Lauer, Damian, & Levelt, 2002).

Even among the relatively small set of objects that are nameable, there is often disagreement within and across languages regarding the most appropriate name for a given object. In developing picture norms and materials for research, name agreement has been considered a critical variable (e.g., Bates et al., 2003; Cuetos et al., 1999; Sanfeliu & Fernandez, 1996; Snodgrass & Vanderwart, 1980; Szekely et al., 2004; Yoon et al., 2004). Ironically, although words are potentially more ambiguous than pictures, most past studies that have examined words in one language and their translation in another language have relied on materials judged subjectively to have only one, or a clearly dominant, translation.

Recently, the question of ambiguity in translation has been raised in several studies (Schonpflug, 1997; Tokowicz & Kroll, 2007; Tokowicz, Prior, & Kroll, 2007). Translation equivalents may have a one-to-many mapping for different reasons. In some cases, within language ambiguity might lead to multiple translations. For example,

A. Prior, aprior@andrew.cmu.edu

the English word *glass* has two distinct meanings: the material and the drinking vessel. Each of these translates onto a different Spanish word: *vidrio* for the former and *vaso* for the latter. Within language, synonymy can also lead to multiple translations: The Spanish word *sofá* may be translated into English as either *sofa* or *couch*. Part of speech, or grammatical class, ambiguity also often results in multiple translations. The English word *cook* can mean either the action (i.e., the verb), in which case it translates into the Spanish *cocinar*, or the person (i.e., the noun), in which case it translates to the Spanish *cocinero*. Finally, there are cases in which multiple translations are a result of the differences in the conceptual–lexical mappings of the two languages. The Spanish noun *reloj* covers the concepts denoted by both *clock* and *watch* in English, each of which is a correct translation. In the same way, the meaning of the English verb *know*, which covers both knowing facts and knowing people, is carried by two distinct verbs in Spanish: *saber* for the former and *conocer* for the latter.

In developing norms for the number of translations for words in Dutch and in English, Tokowicz et al. (2002) examined items used in previous research that had been carefully selected to have only a single or at least a clearly dominant translation. However, of these items, in fact 25% were given more than one translation in each direction of translation. This finding leads to several interesting questions.

An important theoretical issue is the actual prevalence of single-translation items in the bilingual lexicon. If most past research studies used items that apparently had a single translation, whereas in fact most of the translation equivalents in the bilingual lexicon are not unique, then perhaps models developed on the basis of these studies do not adequately represent the varieties of bilingual conceptual and lexical representation. As stated above, Tokowicz et al. (2002) found 25% of items to have more than a single translation, but this is likely to be an underestimation of the actual prevalence of ambiguity in translation because the items sampled in that study had been pre-selected by experimenters to have only a single translation.

Nevertheless, even within their sample, Tokowicz et al. (2002) demonstrated the implications of ambiguity by showing that words that have more than a single translation are judged to be less semantically similar to each of their possible translations than are words that have only a single translation. Furthermore, the number of translations across languages has recently also been shown to affect performance in both recognition and production tasks (e.g., Prior, Kroll, & MacWhinney, 2006; Tokowicz & Kroll, 2007; Tokowicz et al., 2007). Therefore, models of bilingual processing must account for the effects of response competition at various levels.

In this article, we present a set of translation norms for English and Spanish words, based on bilingual participants providing a single written translation for each word presented to them. To our knowledge, this is the first compilation of such norms to be published for these languages. In addition to providing a valuable resource to bilingualism and language production researchers working with this common language pairing, this study addresses

several important issues, outlined below, by sampling a wide range of lexical items in the two languages. We also examine the relations between translation likelihood or probability and a host of psycholinguistic lexical variables that have been previously studied in the monolingual and bilingual literature.

Word Class

As has been noted by both linguists (Levin & Rapaport Hovav, 1996; Wierzbicka, 1988) and psycholinguists (e.g., Gentner, 1981), lexical items belonging to different grammatical classes encode different types of meaning. Nouns typically encode entities, tend to be more perceptually grounded, and their meaning is usually less context-dependent (but see Barsalou, 1982). Verbs, on the other hand, usually encode relations (Ferretti, McRae, & Hatherell, 2001), have more senses (Miller & Fellbaum, 1991), and can be more easily adjusted by contextual demands.

Of special importance for the present discussion, the meanings of verbs and the conceptual aspects encoded in them have greater interlanguage variation than do nominal concepts as demonstrated by Gentner (1981). In the same vein, Van Hell and De Groot (1998) found greater cross-language associative similarity for nouns than for verbs, hinting that nominal translation equivalents share more conceptual features than do verbal concepts.

Despite these word class differences, psycholinguistic bilingual investigations to date have focused almost exclusively on nouns, and with a great emphasis on concrete nouns (e.g., Kroll & Stewart, 1994; La Heij, Hooglander, Kerling, & Van der Velden, 1996; Potter, So, Von Eckardt, & Feldman, 1984). A goal of the present research is to provide more information concerning the generalization of such results to the representation of other grammatical classes, namely verbs. To address this issue, we examine the relation between grammatical class and number of translations. The work of Gentner (1981) and Van Hell and De Groot (1998) suggests that verbs may be more ambiguous in translation than nouns. The present sample included words from both grammatical classes in order to investigate this possibility.

Lexical Properties

The words in our sample were selected to cover a range of printed frequency, imageability, concreteness and age of acquisition (AoA), enabling us to examine whether these within language factors impact the number of different translations a word may have across languages. Tokowicz et al. (2002) reported significant correlations between concreteness and number of translations, such that the concrete nouns in their sample tended to have fewer translations than did the abstract nouns. In the current study we revisit this issue, to see whether it is replicated in our sample. We also examined additional possible predictors for number of translations, including word class as discussed above, word frequency, imageability and AoA.

A second topic of interest concerns those words that can be translated in multiple ways. For these items, we tried to identify factors that might predict translation choice—that is, which of the possible translations is more

likely to be produced. The translation of these words allows the participant the freedom of response choice as long as the meaning of the original word is preserved, and in that sense has some similarity to a within-language free association task. We therefore examined variables known to exert their influence in the free association task, and specifically the frequency of the response word (Nelson, McEvoy, & Dennis, 2000). Free associations given to a certain word tend to have higher frequency than the stimulus word, which might reflect a production bias toward high frequency words, a phenomenon that has been observed in picture and object naming as well (Cuetos, Ellis, & Alvarez, 1999; Snodgrass & Yuditsky, 1996). If this is the case, we would expect to find a similar pattern in translation. Given that for specific multiple translation words there are several possible translations, we set forth to investigate whether the probability of participants providing each of the different translations is related to their frequency in the target language.

A further issue, which is unique for translation generation as opposed to within-language free association, is the question of the form similarity of the translation response to the original word. Translation pairs that overlap in lexical form as well as meaning are considered to be cognates. Cognates typically include a range of similarity across the orthography and phonology of the words in each language. There is a large body of evidence showing that cognate translations are produced more rapidly and accurately than noncognate translations (e.g., De Groot, 1992). Further, when bilinguals perform picture naming exclusively in one of their languages, there is reliable facilitation when the name of the picture in the other language is a cognate (Costa, Caramazza, & Sebastián-Gallés, 2000). In light of these findings, we examined whether in cases where there are several possible translations the probability of giving a specific translation was influenced by its cognate status in relation to the stimulus word. Since cognate translation pairs overlap both in form and in meaning, there might be a bias toward producing a cognate translation, if such a translation exists for a specific item.

Proficiency and Direction of Translation

In this study, we collected translation data from a mixed bilingual population to effectively separate the direction of translation from the specific languages. Thus, English translations for Spanish words were collected from both English-dominant bilinguals (who performed backward translation from their L2 to their L1) and Spanish-dominant bilinguals (who performed forward translation from their L1 into their L2), and the reverse was true for Spanish translations of English words. This choice enables us to isolate language-specific characteristics. For example, word class ambiguity is far more prevalent in English than in Spanish, and in our sample half of the participants encountered its effect in forward translation (from L1 to L2) and the other half encountered its effect in backward translation (from L2 to L1). In addition, the participants in both dominance groups varied in their L2 proficiency, allowing us to examine what role this factor might have on translation choice and translation variation.

METHOD

Participants

Forty Spanish-dominant and 40 English-dominant bilinguals participated in the study, and were paid for their participation. We recruited highly proficient bilinguals—selection criteria included studying the second language for a minimum of 5–6 college semesters or having commensurate language experience. All of the Spanish-dominant bilinguals were immersed in their L2 environment, since the study was conducted in Pittsburgh, for periods ranging from 6 weeks to 34 years, with a median of 4 years. Of the English-dominant bilinguals, 31 had immersion experience in the L2 environment, for periods ranging from 2 weeks up to 20 years, with a median of 7.5 months. In the original sample, there were 7 subjects who reported their first language to have been Spanish but their current dominant language to be English, by virtue of spending long periods of time in the US. We were concerned that these switched-dominance participants might differ from the rest of the English-dominant bilinguals in various aspects, and so replaced them with participants more similar to the rest of the group.

Participants completed a language history questionnaire (LHQ) prior to completing the translation task. Language dominance was assessed as follows: If there were differences in the self-ratings of proficiency in the two languages on the LHQ scales, the language rated by the bilingual as his or her stronger language was assumed to be the dominant language. Participants who rated themselves equally proficient in English and in Spanish were questioned orally, and asked if they had to make a forced choice, which language they would select as being their stronger language. If they were able to make such a choice, their assigned dominance reflected this choice. Finally, those few participants who were unable to make the determination were assumed to be English dominant, by virtue of currently residing in a predominantly English speaking environment.

The LHQ data can be seen in Table 1. The Spanish-dominant bilinguals differed from the English-dominant bilinguals in several aspects: They were older [$t(79) = 5.2, p < .01$]; they had had longer immersion experiences [$t(79) = 4, p < .01$]; less time had passed since their most recent immersion experience, as in fact they were all immersed in their L2 at the time of testing [$t(79) = 2.7, p < .01$]; they had higher self-ratings of L2 proficiency [$t(79) = 4.12, p < .01$]; they reported using their L2 more on a daily basis [$t(79) = 10.4, p < .01$]; and, finally, they had higher performance on the L2 lexical decision task [$t(79) = 6.3, p < .01$] (see details below).

Half of the participants from each dominance group performed English to Spanish translation, and the rest translated in the opposite direction. Importantly, for both dominance groups there were no significant differences between the subgroups that performed each direction of translation (all p values $> .10$).

Materials

Translations were collected first for a set of English words. As a second step, all Spanish translations produced by at least two participants were then normed back and translated by a second group of participants into English. The original set included 670 English words: 241 nouns, 79 verbs, and 350 word-class ambiguous items (e.g., *dress*, which can be both the action and the garment). Words had a frequency of 1–1,290 per million (Kučera & Francis, 1967), with a mean of 109.3 and a SD of 148.9.¹ Ratings of imageability (Bird, Franklin, & Howard, 2001), concreteness (Coltheart, 1981; Wilson, 1988), familiarity (Coltheart, 1981; Wilson, 1988), and AoA (Bird et al., 2001; Coltheart, 1981; Wilson, 1988) were available for a majority of the items in the set.

The Spanish items included 762 words: 525 nouns and 237 verbs. Words had a frequency of 0.5–2,053 per million, with a mean of 82.9 and a SD of 161.3 (Pérez, Alameda, & Cuetos, 2003). Ratings of imageability, concreteness, and familiarity were available for a substantial subset (from LEXESP Sebastián-Gallés, Martí, Cuetos & Carreiras, 2000; using B-pal, Davis & Perea, 2005). AoA ratings were not available for a large enough portion of the sample.

Table 1
Summary of Language History Questionnaire Data

	English-to-Spanish Translation				Spanish-to-English Translation			
	English Dominant		Spanish Dominant		English Dominant		Spanish Dominant	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Age (years)	25.8	9.7	35.9	7.1	22.5	7.3	32.5	10.2
Immersion length (months)	16.7*	31.4	47.7	39.9	9.6	15.3	79.1**	67.8
Time since most recent immersion (months)	19.5	45.5	0	0	21.8	50.3	0	0
L2 proficiency	7.4	1.2	8.3	1.3	7.5	1.4	8.7	0.8
L2 use	3.4	0.8	4.7	0.3	3.1	0.8	4.7	0.3
<i>d'</i>	1.4	0.5	2.0	0.7	1.3	0.5	2.2	0.5

*Excluding a single subject who had been immersed in L2 for 20 years. **Excluding two subjects who had been immersed in L2 for 34 and 28 years, respectively.

Cognate ratings were generated by having a separate group of 30 native English speakers, who did not have any knowledge of Spanish, Portuguese, French, or Italian, perform the translation-elicitation task described by Kroll and Stewart (1994; see also Dufour & Kroll, 1995). Cognate ratings based on translation-elicitation from monolingual speakers have been shown to be comparable to ratings obtained from bilingual speakers (Friel & Kennison, 2001). Participants were presented with a list of words in Spanish, and instructed to guess their translation in English. Each word was presented to 10 participants, so that cognate ratings ranged from 0 (none of the English speakers correctly guessed the translation, due to the fact that there was no cross-language form overlap, such as *muñeca-doll*) to 10 (all the participants correctly guessed the translation, for translation equivalents with highly similar lexical form, such as *concepto-concept*).

Lexical Decision

Two versions of a lexical decision task were developed, one in English and one in Spanish. The procedure for selecting the words was based on that described by Kempe and MacWhinney (1996). Each list included 168 word and 168 orthographically and phonotactically legal nonwords. Participants performed the lexical decision task in the L2. The task took approximately 15 min to complete. *d'* measures of accuracy were computed for each participant, as an added online measure of L2 proficiency (see Table 1).

Procedure

Participants completed the LHQ, followed by the lexical decision task in the L2. They then received a typed booklet containing either English or Spanish words. The English list was divided into two versions, each of 335 words, and each including half of the items from each word-class category. The Spanish stimulus pool was also divided into two lists, each of 381 words, again dividing nouns and verb equally between the versions. The Spanish list was longer, because it was compiled only after the translation data for the English items had been collected, and it included all Spanish words that had been produced by at least two participants. Participants were requested to write down the first translation they could think of for each item on the list. Participants worked at their own pace, and took breaks as necessary.

Scoring

Translations were coded for accuracy using the Larousse Spanish–English and English–Spanish dictionary, and by two native English speakers who are instructors of Spanish in the Department of Modern Languages at Carnegie Mellon University. Each translation was also assigned a grammatical category. Conjugated forms of verbs were accepted as correct responses, and were converted to the infinitive for the purpose of counting the number of different translations. Similarly, plural forms of nouns were accepted as correct responses, and were combined with the singular forms for computing number of translations. Spelling mistakes were also accepted, as long as the

intention of the participant was clear and the mistake did not result in a different word in the language.

RESULTS

Number of Translations

In this analysis we explored the factors that are correlated with a particular word having more than a single translation. The dependent variable was the number of distinct, correct translations that each word had received from different participants (collapsed across both language dominance groups). Since AoA ratings were only available for the English words, we analyzed the English and the Spanish items separately.

Table 2 gives the percent of items from each category that received more than a single translation. For the Spanish words, number of translations ranged from 1 to 7. Verbs were found to have significantly more translations than nouns [$F(1,760) = 12.3$, $MS_e = 14.3$, $p < .001$]. For the English words, number of translations ranged from 1 to 9, and was found to vary significantly by word class [$F(2,667) = 25.1$, $MS_e = 1.6$, $p < .001$]. Planned comparisons demonstrated that verbs again had significantly more translations than did nouns, and the word class ambiguous items had significantly more translations than did the verbs (both p values $< .05$). Finally, English words had significantly more translations on average than Spanish words [$F(1,1430) = 23.9$, $MS_e = 19.9$, $p < .001$], probably due to the large percentage of word class ambiguous items in English (over 50% of the sample).

These significant effects of word class confirm our initial hypothesis that verbs would be more ambiguous in translation than nouns. However, items from the different word classes were not matched on various lexical properties, and these might be confounded with word class. We

Table 2
Percent of Items Given More Than One Translation

Language	Word Class					
	Nouns (N)		Verbs (V)		N/V Ambiguous	
	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>
Spanish ($N = 752$)	45	525	55	237	–	–
English ($N = 670$)	42	241	57	79	69	350

therefore examined several possible predictors of number of translations using hierarchical linear regression, and entered the possibly confounding factors into the equation before examining any residual effects of word class. The following analyses examined only the lexical properties of the word in the source language. The properties of the given translation and their influence on translation choice are examined in the following section.

In the analysis of English words, the following variables were entered into the model. Word length and log frequency were entered on the first step and were found to be significant predictors of number of translations [$\Delta R^2 = .029$; $F(2,629) = 9.1$, $p < .001$]. Both factors were negatively correlated with number of translations, such that shorter words were in general more ambiguous in translation, while highly frequent words tended to have fewer translations. The possible contributions of imageability, concreteness, familiarity and AoA as predictors of number of translations were examined next. Since these four lexical properties tended to be highly correlated with each other (R s ranging from 0.53 to 0.89, all p values $< .001$), the procedure suggested by Cohen et al. (2003) was used and identified imageability as the best predictor of number of translations. Therefore, imageability ratings were entered on the second step and were found to significantly predict number of translations [$\Delta R^2 = .028$; $F(1,628) = 18.3$, $p < .001$], such that highly imageable words tended to have fewer translations.

On the third and last step of the analysis, part of speech (POS) was entered into the model and was found to account for a significant portion of the variance in number of translations, even after length, frequency and imageability were controlled for [$\Delta R^2 = .051$; $F(1,627) = 35.9$, $p < .001$]. This reflected the fact that nouns had fewer translations than did verbs, and verbs in turn had fewer translations than did word class ambiguous items.

We then repeated the described analysis separately for number of translations data computed for each of the dominance groups. Thus, the English-dominant bilinguals were translating from L1 to L2, while the Spanish-dominant bilinguals were translating from L2 to L1. In both cases, the same pattern of results was preserved. We found significant influences of word length and log frequency in both cases. Also, for both dominance groups imageability ratings were found to be a superior predictor, when compared with concreteness, familiarity and AoA ratings. Of specific interest, the predictive power of AoA ratings did not differ for English and for Spanish-dominant bilinguals, despite the possible hypothesis that these would have greater influence for the English-dominant bilinguals.² In both cases, the contribution of the AoA ratings to the model was not significant, and did not survive the influence of imageability and concreteness (Cohen et al., 2003). Finally, for both dominance groups we found significant effects of POS, equivalent to those found for the number of translations data examined when collapsing across the groups.

In the analysis of Spanish words, we used the length, log frequency, imageability, and part of speech as predictor variables, based on those used in the previous analysis. As in the analysis of English words, length and log fre-

quency were entered on the first step, and were found to be significant predictors of number of translations [$\Delta R^2 = .016$; $F(2,670) = 5.5$, $p < .01$]. However, an examination of the regression coefficients showed that in fact only word frequency was playing a significant role, such that high frequency words tended to have fewer translations. Imageability was entered on the second step, and was found to significantly predict number of translations [$\Delta R^2 = .021$; $F(1,669) = 14.5$, $p < .01$]. Part of speech was entered on the third step, and did not contribute significantly to the model [$\Delta R^2 = .003$; $F(1,668) = 2.0$, $p > .15$]. This last finding seems to suggest that for our Spanish word sample, the verbs were not in general more ambiguous in translation than the nouns, after controlling for possible word class differences in frequency and imageability. However, this finding might be due to a sampling error, since we did not have imageability ratings available for some of the more ambiguous Spanish verbs, which were therefore not included in the analysis. Thus, for verbs entered in the analysis the average number of translations was 2.06, and for those not entered (due to lack of imageability ratings) the average number of translations was 2.35. The average numbers of translations for nouns entered and not entered in the analyses were virtually identical (1.73 and 1.76, respectively). We are therefore hesitant to conclude that Part of speech does not play a significant role in translation ambiguity in Spanish. This issue requires further examination.

Translation Choice

In the following set of analyses, we investigated the factors that influence a bilingual's choice of any given translation over another when there is more than one possible translation. Thus, the influence of different lexical properties of the possible translations was examined.

In the present set of analyses, translation probability was used as the dependent variable. For each stimulus word that received more than a single translation the probability of each of the responses was calculated as follows: the number of bilinguals giving each response divided by the total number of correct responses given to the word. Furthermore, only words that were ambiguous in translation were entered into the analysis, because these are the only items for which the bilingual must make a choice. The two language pairings were again examined separately.

We wished to determine whether words that are frequent in the target language are given more often as translations. A second factor examined was the form similarity between the cue and target words, namely those cases where the cue word has a form-related cognate translation in the target language. Is the cognate translation chosen more frequently over other possible translations? Finally, the relation of target imageability and translation probability was also examined.

In the analysis of Spanish-to-English translation probability, target word length and log frequency were entered on the first step of the regression, and were found to marginally predict translation probability [$\Delta R^2 = .012$; $F(2,499) = 2.9$, $p = .052$], such that shorter and more frequent words tended to have higher probabilities of being produced as translations. Target imageability was entered

on the second step and significantly predicted translation probability [$\Delta R^2 = .015$; $F(1,498) = 7.8$, $p < .01$], as more imageable words had a higher likelihood of being given as translations. Cognate ratings were entered on the third step and were also found to be a significant and strong predictor of translation probability [$\Delta R^2 = .11$; $F(1,497) = 64.3$, $p < .001$], since words high in form similarity to the cue word were given as translations more often. The two-way interactions were added to the model on the fourth step, and were found to add to it significantly [$\Delta R^2 = .022$; $F(3,494) = 4.34$, $p < .01$]. An examination of the coefficients showed that this was due to the significant interaction between log frequency and cognate rating. The estimated means demonstrating the nature of the interaction can be seen in Figure 1. Essentially, cognate effects were more pronounced for high than for low frequency items. Finally, the three-way interaction was entered on the fifth step of the model and did not have a significant contribution ($\Delta R^2 = 0$).

The analysis of translation from English to Spanish yielded similar patterns. Target word length and log frequency were entered on the first step of the regression and were found to significantly predict translation probability [$\Delta R^2 = .012$; $F(2,526) = 3.1$, $p < .05$]. Target imageability was entered on the second step, and significantly predicted translation probability [$\Delta R^2 = .015$; $F(1,525) = 8.2$, $p < .01$], as more imageable words had a higher likelihood of being given as translations. Cognate ratings were entered on the third step and significantly predicted translation probability [$\Delta R^2 = .052$; $F(1,524) = 29.6$, $p < .001$], since target words high in form similarity to the cue word were more likely to be given as translations. The two-way interactions were added to the model on the fourth step, and did not add to it significantly [$\Delta R^2 = .009$; $F(1,521) = 1.7$, $p > .15$], though the interaction between log frequency and cognate rating was marginally significant ($p = .06$), with a pattern similar to that appearing for the Spanish-to-English translation. Finally, the three-way interaction was added on the fifth step, and did not add significantly to the model [$\Delta R^2 = .00$; $F(1,520) < 1$].

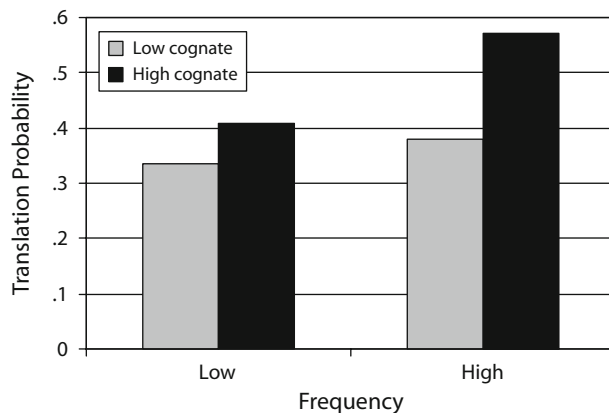


Figure 1. Estimated means illustrating frequency and cognate effects in translation probability for Spanish-to-English translation.

Language Dominance and L2 Proficiency

In this set of analyses, we explored possible differences between the two directions of translation, both in terms of the languages involved and in terms of the language dominance of the bilinguals participating in the study. Thus, we had both Spanish- and English-dominant bilinguals translating from English to Spanish and vice versa. This enabled us to identify patterns that are language specific. Second, we broadened our investigation of translation choice, and tried to identify influences of direction of translation (forward vs. backward) and L2 proficiency in negotiating translation ambiguity.

To address these issues, we calculated how many different translations each item had received, divided by the total number of correct responses. We calculated this value separately for each of the language groups: English dominant and Spanish dominant. We then conducted a two-way repeated measures ANOVA with number of translations as the dependent variable, language as the between items independent variable and dominance group as a within item repeated measure independent variable.

There was a significant main effect of language, showing that across dominance groups English words were more ambiguous in translation than were Spanish words [$F(1,1430) = 24.3$, $p < .001$]. There was also a significant main effect of dominance group [$F(1,1430) = 135$, $p < .001$] since English-dominant bilinguals produced a greater number of distinct translations for the stimulus words in both languages when compared with the Spanish-dominant bilinguals. The two-way interaction was not significant ($p = .3$).

We attribute the main effect of language to the fact that the English language includes many class ambiguous words, a phenomenon that is virtually nonexistent in Spanish. Thus, it is not surprising to find higher ambiguity in translation over all when translating from English to Spanish than in the opposite direction, regardless of whether a specific bilingual is performing forward (L1 to L2) or backward (L2 to L1) translation.

We were at first puzzled by the main effect of language dominance group, namely that the Spanish-dominant bilinguals seemed to be in better agreement with one another when translating items in both their L1 and L2. However, an examination of the LHQ data (see Table 1) suggests that this might reflect the difference in L2 proficiency between the two bilingual groups. In general, the Spanish-dominant bilinguals had higher L2 proficiency (in English) than the English-dominant bilinguals (in Spanish), even though there is a range of proficiency in both groups (see detailed analyses in the participants section).

In order to further test this hypothesis, we calculated for each participant the average probability of the translations he or she produced when translating ambiguous items. Whereas a specific bilingual might almost always produce the dominant translation for ambiguous items, conforming to the popular choice within the group, a different participant might be more likely to produce less probable yet still correct translations, diverging from the selection made by the majority of the group. Thus, the average probability of the translations a bilingual chose

for all ambiguous items reflects her tendency to choose either highly probable or less probable translations when a choice is possible.

We then asked whether these average probability scores were correlated with L2 proficiency measures—both subjective (L2 proficiency self rating; self reporting of L2 daily use) and objective (length of immersion, performance on an L2 lexical decision task). We performed this analysis independently for forward and backward translation, while collapsing across language dominance groups. When our sample is separated by direction of translation (forward vs. backward), the two resulting groups do not differ significantly on any of the measures reported in the previous section: L2 proficiency, L2 use, immersion length, or L2 lexical decision performance (all *p* values > 0.4).

The correlation matrix between the average probability score and the various proficiency measures can be found in Table 3. We found significant correlations between participant average probability scores and proficiency measures in forward translation. These correlations suggest that less proficient L2 speakers tend to have lower average probability scores. Phrased differently, less proficient bilinguals tend to vary more in their selection of a translation for ambiguous items. Thus, as a group, the less proficient bilinguals converge to a lesser degree than more proficient bilinguals onto the “best” translation when translating into their L2.

On the other hand, Table 3 shows no significant correlations between average probability scores and L2 proficiency in backward translation (whereas the proficiency measures continue to be significantly correlated among themselves). Thus, L2 proficiency does not seem to influence backward translation in a manner similar to its influence on forward translation. Similar findings have been reported by Kroll et al. (2002) in a study comparing language learners and proficient bilinguals, which demonstrated that forward translation is much more affected by proficiency than backward translation.

To further investigate this intriguing pattern, we recoded the number of translations data in terms of forward versus backward translation, as opposed to the previous analysis which was coded by the language of the stimulus word (English or Spanish). We then repeated the two-way ANOVA with number of translations as the dependent variable and language dominance and direction of translation as the independent variables. We once again find a main effect of language dominance [$F(1,76) = 16.1, p < .001$], but critically, this was qualified by a significant interaction between language dominance and direction of translation [$F(1,76) = 54.9, p < .001$], indicating that the substantial differences between the English-dominant and the Spanish-dominant groups were evident mainly in forward translation (from L1 to L2), and virtually disappeared in backward translation (see Figure 2). These results align with the outcome of the correlation analysis, and lend support to our previous conclusion, namely that lower L2 proficiency negatively impacts the ability of bilinguals to select high-probability translations in cases where translation ambiguity exists, thus generating a pattern of behavior that is markedly different than that found for more proficient bilingual speakers.

Table 3
Correlation Matrix of Average Translation Probability and L2 Proficiency Measures, by Direction of Translation

	1	2	3	4
L1 to L2				
1. Average probability	1			
2. Immersion length	.332*	1		
3. L2 proficiency	.509**	.380*	1	
4. L2 use	.691**	.387*	.610**	1
5. L2 lexical decision <i>d'</i>	.608**	.431**	.568**	.376*
L2 to L1				
1. Average probability	1			
2. Immersion length	n.s.	1		
3. L2 proficiency	n.s.	.301†	1	
4. L2 use	n.s.	.389*	.254†	1
5. L2 lexical decision <i>d'</i>	n.s.	.416**	.447**	.474**

p* < .05. *p* < .01. †*p* < .1.

DISCUSSION

In this article, we present a set of translation norms for large samples of English and Spanish words. The norms include nouns, verbs, and class-ambiguous words covering a wide range of frequency, imageability and cross-linguistic form overlap. These norms will hopefully serve as a useful tool for investigators of bilingual language processing and language production in general.

Like the results of Tokowicz et al. (2002) for Dutch and English, these data for Spanish and English show that translation ambiguity is highly prevalent. Almost 50% of the Spanish words and 60% of the English words were found to have more than a single translation. These numbers alone highlight the importance of translation ambiguity in two ways. First, ambiguity must be carefully considered and controlled in the construction of experimental materials. Second, translation ambiguity, its role in bilingual representation and performance and its interaction with other established lexical variables merit investigation in their own right. Indeed, recent findings demonstrate the role played by ambiguity and competition in translation performance (Tokowicz & Kroll, 2007; Tokowicz et al.,

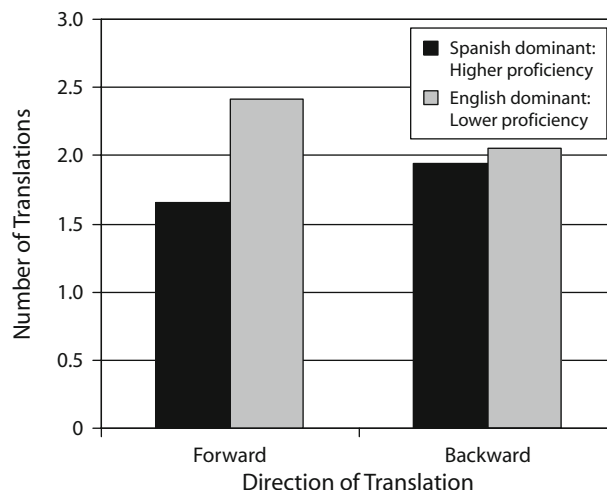


Figure 2. Number of translations by direction of translation.

2007; Prior et al., 2006). Finally, we found that languages may differ in their degree of ambiguity (see also Bates et al., 2003). Specifically in the present case English was found to be more ambiguous in translation than Spanish, due to widespread word-class ambiguity in English. Similar investigations of additional languages will provide more data regarding these cross-linguistic differences, and their possible implications for language learning.

The present investigation goes further in demonstrating that ambiguity in translation is not purely accidental. Thus, word frequency and imageability are correlated with the likelihood of a word being ambiguous in translation in a predictable manner. Most notably, we found significant differences in the prevalence of ambiguity for words belonging to different grammatical classes. Verbs (and class ambiguous words) tended to be more ambiguous in translation than nouns. These results are commensurate with previous research suggesting word class differences in the degree of cross-linguistic meaning overlap (Gentner, 1981; Van Hell & De Groot, 1998). Furthermore, they underline the importance of extending the scope of investigations of bilingual representation and processing to include verbs. Most research to date has focused on nouns, and has identified several lexical variables that might play a role in the crosslinguistic representation of meaning. For instance, in the distributed feature model, De Groot (1992) suggests that concrete nouns and cognates might share more meaning features between the two languages of a bilingual speaker than abstract nouns and noncognates. Future research should examine the degree to which previous findings generated using nominal materials can be generalized to other word-classes (see Prior et al., 2006, for preliminary findings).

Given that translation ambiguity is so prevalent, the question arises how bilinguals negotiate the competition arising between possible translations. The present study examined offline performance only, and thus cannot speak to the temporal dynamics of competition resolution (for evidence on the consequences for processing see Prior et al., 2006; Tokowicz & Kroll, 2007; Tokowicz et al., 2007). Nevertheless, we were able to identify lexical variables that influence the outcome and allow us to gain insight into what types of translations bilinguals tend to prefer when given the choice. Bilingual speakers were found to rely on lexical characteristics of the possible translations in the target language, gravitating toward more frequent words, and words that were more imageable. Bilinguals were also influenced by the degree of cross-linguistic form overlap of the possible translations with the stimulus words, showing a clear preference for a cognate translation, if one existed. These variables have been identified in past research as facilitating translation performance, and leading to faster and more accurate translation of nonambiguous words (De Groot, 1992). Further, word frequency is also known to play an important role in other language production contexts, such as picture naming (e.g., Cuetos, Ellis, & Alvarez, 1999) and free association (Nelson, McEvoy, & Dennis, 2000). Therefore, to the degree that translation is conceived as a language production task (albeit written production in the present case) it stands to reason that similar variables will exert their influence.

A final focus of this study concerned the relation between L2 proficiency and effectively negotiating translation ambiguity. Our findings suggest that in forward translation, which requires production in the L2, a bilingual's second language proficiency can influence her ability to successfully align her translation choice with that made by other bilinguals. Specifically, less proficient bilinguals were found to produce low-probability translations more often than more proficient bilinguals. This disparity might cause less proficient speakers to produce nonnative sounding speech and might also negatively impact their ability to comprehend rapid input, due to activation of forms less likely to be encountered.

The fact that lower proficiency bilinguals produced more low-probability translations might be due to retrieval difficulties, causing the less proficient bilingual to settle for a less likely (though still correct) translation due to a momentary inability to produce the higher probability translation. Alternatively, this pattern might be a result of knowledge gaps, in the sense that lower proficiency bilinguals might simply not know the higher probability translation of a specific word, and might be forced to produce the only translation available to them.

A third possibility is that, although the less proficient bilinguals in our sample possessed the lexical knowledge of all the translation alternatives, and were not hindered by temporary retrieval difficulties, their lexical or conceptual representations differ from those of more proficient speakers of the language. These disparities might lead less proficient bilinguals to choose translations that diverge from the norm set by more proficient speakers. With improving proficiency, bilinguals are better able to converge onto and agree upon the likelihood of specific translations, which reflects their growing knowledge and fine-tuning of meaning and lexical representations within the L2 (see also Malt & Sloman, 2003; Zhang, 1995).

The data reported here do not allow us to prefer one of these accounts over the other. Moreover, these influences are not mutually exclusive, and might operate simultaneously. Further research will be needed to better understand the specific mechanisms at play.

To conclude, the present study provides translation norms for a large set of English and Spanish words. These norms can be used by researchers interested in bilingual and monolingual language performance. The results demonstrate systematic relations between translation ambiguity on the one hand, and word frequency, word imageability and word class on the other hand. We also found that these same lexical properties, as well as cognate ratings, can predict translation choice in cases where ambiguity exists. Finally, L2 proficiency was found to play a significant role in negotiating translation ambiguity, at least when translating into the L2. Jointly, these findings highlight the importance of translation ambiguity as a factor influencing bilingual representation and performance.

AUTHOR NOTE

This research was supported by NICHD Postdoctoral NRSA Award F32HD049255 to A.P. and by NSF Grant BCS-0418071 to J.F.K. We thank Tamar Degani, Natasha Tokowicz, and an anonymous reviewer

for helpful comments; Frances Ruiz and Therese Tardio for assistance in coding translation accuracy; and Mercedes Farrell, Anna Guitchouts, and Tyler Phelps for diligent data collection. Correspondence concerning this article should be addressed to A. Prior, Department of Psychology, 254P Baker Hall, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15217 (e-mail: aprior@andrew.cmu.edu).

REFERENCES

- BARSALOU, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition*, **10**, 82-93.
- BATES, E., D'AMICO, S., JACOBSEN, T., SZÉKELY, A., ANDONOVA, E., DEVESCOVI, A., ET AL. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, **10**, 344-380.
- BIRD, H., FRANKLIN, S., & HOWARD, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers*, **33**, 73-79.
- CHEN, H. C., & NG, M. L. (1989). Semantic facilitation and translation priming effects in Chinese-English bilinguals. *Memory & Cognition*, **17**, 454-462.
- COHEN, J., COHEN, P., WEST, S. G., & AIKEN, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: Erlbaum.
- COLTHEART, M. (1981). *MRC psycholinguistic database: User manual Version 1*. Retrieved in November 2006 from www.psych.rl.ac.uk/User_Manual_v1_0.html.
- COSTA, A., CARAMAZZA, A., & SEBASTIÁN-GALLÉS, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 1283-1296.
- COSTA, A., MIOZZO, M., & CARAMAZZA, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory & Language*, **41**, 365-397.
- CUETOS, F., ELLIS, A. W., & ALVAREZ, B. (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments, & Computers*, **31**, 650-658.
- DAVIS, C. J., & PEREA, M. (2005). BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods*, **37**, 665-671.
- DE GROOT, A. M. B. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 1001-1018.
- DUFOUR, R., & KROLL, J. F. (1995). Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation. *Memory & Cognition*, **23**, 166-180.
- FERRETTI, T. R., MCRAE, K., & HATHERELL, A. (2001). Integrating verbs, situation schemas and thematic role concepts. *Journal of Memory & Language*, **44**, 516-547.
- FRIEL, B. M., & KENNISON, S. M. (2001). Identifying German-English cognates, false cognates and non-cognates: Methodological issues and descriptive norms. *Bilingualism: Language & Cognition*, **4**, 249-274.
- GENTNER, D. (1981). Some interesting differences between verbs and nouns. *Cognition & Brain Theory*, **4**, 161-177.
- GOLLAN, T., FORSTER, K. I., & FROST, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 1122-1139.
- GUMNIOR, J., BOLTE, J., & ZWITZERLOOD, P. (2006). A chatterbox is a box: Morphology in German word production. *Language & Cognitive Processes*, **21**, 920-944.
- HERMANS, D., BONGAERTS, T., DE BOT, K., & SCHREUDER, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language & Cognition*, **1**, 213-229.
- JESCHENIAK, J. D., & LEVELT, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and phonological form. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 824-843.
- JIANG, N. (2000). Lexical representation and development in a second language. *Applied Linguistics*, **21**, 47-77.
- KEATLEY, C., SPINKS, J., & DE GELDER, B. (1994). Asymmetrical semantic facilitation between languages. *Memory & Cognition*, **22**, 70-84.
- KEMPE, V., & MACWHINNEY, B. (1996). The crosslinguistic assessment of foreign language vocabulary learning. *Applied Psycholinguistics*, **17**, 149-183.
- KROLL, J. F., MICHEAL, E., TOKOWICZ, T., & DUFOUR, R. (2002). The development of lexical fluency in a second language. *Second Language Research*, **18**, 137-171.
- KROLL, J. F., & STEWART, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory & Language*, **33**, 149-174.
- KROLL, J. F., & TOKOWICZ, N. (2005). Models of bilingual representation and processing: Looking back and to the future. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 531-553). New York: Oxford University Press.
- KUČERA, H., & FRANCIS, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- LA HEIJ, W., HOOGLANDER, A., KERLING, R., & VAN DER VELDEN, E. (1996). Nonverbal context effects in forward and backward word translation: Evidence for concept mediation. *Journal of Memory & Language*, **35**, 648-665.
- LEVIN, B., & RAPPAPORT HOVAV, M. (1996). *From lexical semantics to argument realization*. Unpublished manuscript, Northwestern University, Evanston, IL, and Bar Ilan University, Ramat Gan, Israel.
- MALT, B. C., & SLOMAN, S. A. (2003). Linguistic diversity and object naming by non-native speakers of English. *Bilingualism: Language & Cognition*, **6**, 47-67.
- MILLER, G. A., & FELLBAUM, C. (1991). Semantic networks of English. *Cognition*, **41**, 197-229.
- NELSON, D. L., MCEVOY, C. L., & DENNIS, S. (2000). What is free association and what does it measure? *Memory & Cognition*, **28**, 887-899.
- PÉREZ, M. A., ALAMEDA, J. R., & CUETOS, F. (2003). Frecuencia, longitud y vecindad ortográfica de las palabras de 3 a 16 letras del diccionario de la lengua Española (RAE, 1992). *Revista Electrónica de Metodología Aplicada*, **8**, 1-10.
- POTTER, M., SO, K., VON ECKARDT, B., & FELDMAN, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning & Verbal Behavior*, **23**, 23-38.
- PRIOR, A., KROLL, J. F., & MACWHINNEY, B. (2006, November). *The role of translation probability and word class in two translation tasks*. Poster presented at the 47th Annual meeting of the Psychonomic Society, Houston.
- 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 & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 226-250). New York: Oxford University Press.
- SANFELIU, M. C., & FERNANDEZ, A. (1996). A set of 254 Snodgrass-Vanderwart pictures standardized for Spanish: Norms for name agreement, image agreement, familiarity, and visual complexity. *Behavior Research Methods, Instruments, & Computers*, **28**, 537-555.
- SCHONPFLUG, U. (1997). *Bilingualism and memory*. Paper presented at the International Symposium on Bilingualism, Newcastle-Upon-Tyne, U.K.
- SEBASTIÁN-GALLÉS, N., MARTÍ, M. A., CUETOS, F., & CARREIRAS, M. (2000). *LEXESP: Léxico informatizado del español*. Barcelona: Edicions de la Universitat de Barcelona.
- SNODGRASS, J. G., & VANDERWART, M. A. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **6**, 174-215.
- SNODGRASS, J. G., & YUDITSKY, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, & Computers*, **28**, 516-536.
- SZEKELY, A., JACOBSEN, T., D'AMICO, S., DEVESCOVI, A., ANDONOVA, E., HERRON, D., ET AL. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory & Language*, **51**, 247-250.
- TOKOWICZ, N., & KROLL, J. F. (2007). Number of meanings and concreteness: Consequences of ambiguity within and across languages. *Language & Cognitive Processes*, **22**, 727-779.

- TOKOWICZ, N., KROLL, J. F., DE GROOT, A. M. B., & VAN HELL, J. G. (2002). Number-of-translation norms for Dutch–English translation pairs: A new tool for examining language production. *Behavior Research Methods, Instruments, & Computers*, **34**, 435–451.
- TOKOWICZ, N., PRIOR, A., & KROLL, J. F. (2007). *The role of translation ambiguity and semantic similarity in translation production*. Manuscript in preparation.
- VAN HELL, J. G., & DE GROOT, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language & Cognition*, **1**, 193–211.
- VIGLIOTTO, G., LAUER, M., DAMIAN, M. F., & LEVELT, W. J. M. (2002). Semantic and syntactic forces in noun phrase production. *Journal of Experimental Psychology: Learning, Memory & Cognition*, **28**, 46–58.
- WIERZBICKA, A. (1988). *The semantics of grammar*. Amsterdam: Benjamins.
- WILSON, M. D. (1988). The MRC Psycholinguistic Database: Machine readable dictionary, Version 2. *Behavior Research Methods, Instruments, & Computers*, **20**, 6–11.
- YOON, C., FEINBERG, F., LUO, T., HEDDEN, T., GUTCHESS, A. H., CHEN, H.-Y. M., ET AL. (2004). A cross-culturally standardized set of pictures for younger and older adults: American and Chinese norms for name agreement, concept agreement, and familiarity. *Behavior Research Methods, Instruments, & Computers*, **36**, 639–549.
- ZHANG, S. (1995). Semantic differentiation in the acquisition of English as a second language. *Language Learning*, **45**, 225–249.

NOTES

1. The three most frequent items (*do*, 1,363; *have*, 3,941; *be*, 6377) were excluded from these calculations.
2. This is because the AoA ratings we used were for the acquisition of English as a first language.

ARCHIVED MATERIALS

The following materials may be accessed through the Psychonomic Society's Norms, Stimuli, and Data Archive, www.psychonomic.org/Archive.

To access these files or links, search the archive for this article using the journal (*Behavior Research Methods*), the first author's name (Prior), and the publication year (2007).

FILE: Prior-BRM-2007.zip

DESCRIPTION: The compressed archive file contains nine files:

EnglishToSpanish_master.xls contains each English cue word once and all the different translations it received are listed on the same row, in decreasing order of probability.

SpanishToEnglish_Master.xls contains each Spanish cue word once and all the different translations it received are listed on the same row, in decreasing order of probability.

EnglishToSpanish_TranslationPairs.xls contains each unique pair of English cue–Spanish translation listed on a separate line, and includes lexical variables of the cue and translation words.

SpanishToEnglish_TranslationPairs.xls contains each unique pair of Spanish cue–English translation listed on a separate line, and includes lexical variables of the cue and translation words.

Readme.txt gives full details of the columns included in each file.

The compressed archive file also includes .csv (comma-separated values) versions of the first four files, which were created using Microsoft Excel on Windows. However, some of the unique Spanish characters might be corrupted in these files.

AUTHOR'S E-MAIL ADDRESS: aprior@andrew.cmu.edu.

AUTHOR'S WEB SITE: www.andrew.cmu.edu/user/aprior/

(Manuscript received January 12, 2007;
revision accepted for publication February 12, 2007.)