Published in final edited form as:

Linguist Approaches Biling. 2016 ; 6(3): 290–307. doi:10.1075/lab.14013.deg.

Determinants of translation ambiguity:

A within and cross-language comparison

Tamar Degani[1], **Anat Prior**[1], **Chelsea M. Eddington**[2], **Ana B. Arêas da Luz Fontes**[3], and **Natasha Tokowicz**[2]

[1]University of Haifa

[2]University of Pittsburgh

[3]Universidade Federal do Rio Grande do Sul

Abstract

Ambiguity in translation is highly prevalent, and has consequences for second-language learning and for bilingual lexical processing. To better understand this phenomenon, the current study compared the determinants of translation ambiguity across four sets of translation norms from English to Spanish, Dutch, German and Hebrew. The number of translations an English word received was correlated across these different languages, and was also correlated with the number of senses the word has in English, demonstrating that translation ambiguity is partially determined by within-language semantic ambiguity. For semantically-ambiguous English words, the probability of the different translations in Spanish and Hebrew was predicted by the meaning-dominance structure in English, beyond the influence of other lexical and semantic factors, for bilinguals translating from their L1, and translating from their L2. These findings are consistent with models postulating direct access to meaning from L2 words for moderately-proficient bilinguals.

1. Introduction

Bilinguals are often confronted with the situation in which translation equivalents do not align in a one-to-one fashion across languages, and instead exhibit ‘translation ambiguity’ in that more than one translation is possible for a given word (e.g., Tokowicz, Kroll, de Groot, & van Hell, 2002). For instance, the word ‘proposal’ can be translated into Dutch as

Authors’ addresses
Tamar Degani, Department of Communication Sciences and Disorders University of Haifa, Mount Carmel, 199 Aba Khoushy Ave, Haifa, 3498838, ISRAEL
Anat Prior, Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, University of Haifa, ISRAEL
Chelsea M. Eddington, Department of Psychology and the Center for the Neural Basis of Cognition, University of Pittsburgh, Pittsburgh, PA, USA
Ana B. Arêas da Luz Fontes, Professora Adjunta de Línguas Modernas, Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, BRASIL
Natasha Tokowicz, Departments of Psychology and Linguistics and the Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA, USA

A version of this work was presented at the 50th Annual Meeting of the Psychonomic Society, in November 2009.

‘ voorstel’ or ‘aanzoek’. Such translation ambiguity can be a result of ambiguity within the source language, or of target language characteristics such as synonymy. In the current study, we compare four sets of translation norms, documenting translation ambiguity between English (the shared source language across the four norms) and Dutch, German, Spanish, and Hebrew as target languages. This comparison allows us to identify to what degree the phenomenon of translation ambiguity is driven by source language factors, with a specific emphasis on within-language meaning ambiguity. To the extent that ambiguity in the source language drives translation ambiguity, we would expect a high degree of correspondence in the number of translations that an English word receives in different languages. Characterizing the phenomenon of translation ambiguity is important because previous studies have demonstrated that translation ambiguity has consequences for bilingual language processing (Boada, Sanchez-Casas, Gavilan, Garcia-Albea, & Tokowicz, 2013) and learning (Degani & Tokowicz, 2010) and that it may be confounded with other relevant factors such as concreteness (Tokowicz & Kroll, 2007) and part-of-speech (Prior, Kroll, & MacWhinney, 2013).

1.1 Translation ambiguity – prevalence and potential sources

The degree of ambiguity in translation of a specific word can be assessed in different manners. One straightforward approach would be to count dictionary entries in a bilingual dictionary. However, this method would arguably over-estimate the prevalence of translation ambiguity, and further might not yield a valid reflection of the actual lexical knowledge of bilingual speakers, who might not know all dictionary entries (Tokowicz et al., 2002). In the current study, therefore, we focus on the actual knowledge of bilinguals and adopt the ‘first translation’ method. In this approach, bilinguals are requested to translate single words from the source to the target language, giving only the first translation that comes to mind. Translation ambiguity is consequently assessed as the number of different correct translations a given word receives across the bilingual participants. Further, in cases in which the sample of responders is large enough, the probability of the various translations can also be estimated. Notably, because a single translation is provided by each participant, one concern could be that only high-frequency translations would be provided, leading to some underestimation of translation ambiguity. However, the first translation methodology has been adopted from within-language free association studies (e.g., Twilley, Dixon, Taylor, & Clark, 1994). In such an approach, the stochastic properties of the lexical representation network allow for less frequently used options to be produced occasionally, as is the case in monolingual naming studies (for example, Peterson & Savoy, 1998). Indeed, in one study that examined this issue, although lexical frequency in the target language was found to influence translation choice, its influence was rather weak (Prior, MacWhinney, & Kroll, 2007).

The translation norms used in the current study have all relied on the ‘first translation’ method, and all include English as the source language. Thus, in the current study we are able to compare translation ambiguity across several language pairs to arrive at a better understanding of the underpinnings of this phenomenon.

In the first published investigation of translation ambiguity, Tokowicz et al. (2002) collected number of translation norms from Dutch-English proficient bilinguals. They found that approximately 25% of 562 words were translation ambiguous in that they elicited more than one translation across participants in one or both directions of translation. Further, 40% of these same English words elicited more than one translation in German (Eddington, Degani, & Tokowicz, 2014). These estimates are likely somewhat of an underestimation, however, because the stimuli in these studies were chosen from previous research on bilingual processing in which they were assumed to have a single translation across languages. Indeed, Prior et al. (2007) investigated a different set of 670 English words, and observed that about 60% of the items received multiple translations to Spanish. Finally, roughly 55% of these same items received multiple translations across Hebrew and English (Smith, Walters, & Prior, 2012). In the current study, we used these four sets of norms to examine factors contributing to cross-language translation ambiguity.

Several factors pertaining to the source language as well as the target language have been discussed as potential causes of translation ambiguity (see e.g., Prior, Wintner, MacWhinney, & Lavie, 2011). First, and most relevant to the current investigation, semantic ambiguity in the source language may lead to the existence of multiple translations, because each meaning is likely to receive a different translation in a different language (Frenck-Mestre & Prince, 1997; but see Degani & Tokowicz, 2013). For example, each meaning of the English word ‘mean’ may receive a different translation in another language (in Hebrew, for example, ‘rasha’ for *evil* and ‘memutza’ for the *average* meaning). To the extent that semantic ambiguity in the source language underlies translation ambiguity, we would expect a high degree of correlation in number of translations from English to Dutch, English to German, English to Spanish, and English to Hebrew. Moreover, the relative dominance of each meaning in the source language might impact the probability with which the corresponding translation is given in a translation task. For example, because the evil meaning of ‘mean’ in English is more frequent (Twilley et al., 1994), we would expect its translation ‘rasha’ to be given more frequently in a translation task from English to Hebrew.

In some cases, a word in a given language may capture multiple related senses. Such polysemy may carry over to another language, such that a single word in the target language also captures these same multiple related senses (Degani & Tokowicz, 2013). For example, the two senses of the English word ‘market’ (flea market and housing market) are captured by the single Spanish translation ‘mercado’. Critically, however, in some cases the target language specifies unique lexical labels for each sense of the source language word. For example, the English verb ‘to know’ is translated to Hebrew as ‘ladaat’ when relating to factual knowledge and as ‘lehakir’ when relating to knowing a person (the same differentiation also occurs in Spanish, with the translations ‘saber’ and ‘conocer’ respectively). To the degree that certain source language lexical items denote wide conceptual spaces, other languages may divide the conceptual space more finely, resulting in more than a single translation for such words. This would further contribute to correspondence in the number of translations across the various language pairs.

Second, part-of-speech and morphological ambiguity in the source language are also likely to elicit multiple translations. Prior et al. (2007) showed that words that are part-of-speech

ambiguous in English tend to receive multiple translations, each denoting a different word class (e.g., ‘cocinero’ in Spanish to denote the noun meaning, and ‘cocinar’ to denote the verb meaning of the English word ‘cook’). Correlations among number of translations across different languages may also be the result of such part-of-speech and morphological ambiguity in the source language.

Finally, characteristics of the target language, such as near synonymy, can also lead to the availability of more than one translation for a given word. For instance, when translating the English word *car* into Hebrew, both ‘otto’ and ‘mexonit’ serve as correct alternatives for translation. In contrast to source language factors discussed above, such target language synonymy is likely to reduce the correlations among the different number of translation norms, because it reflects distinctive properties of each target language.

1.2 Effects of translation ambiguity on bilingual performance

Translation ambiguity is relevant not only because it is widespread (Eddington et al., 2014; Prior et al., 2007; Tokowicz et al., 2002), but also because it affects learning and processing of bilingual speakers across a wide range of proficiencies (for review, see Tokowicz & Degani, 2010). Degani and Tokowicz (2010) showed that learning words that are translation ambiguous is more difficult than learning unambiguous words. Native English speakers had more difficulty learning Dutch words that shared an English translation (e.g., two Dutch synonyms for the word ‘boot’) compared to Dutch words that were learned with unique English translations (see also Degani, Tseng, & Tokowicz, 2014).

The influence of translation ambiguity is not limited, however, to beginning L2 learners. Tokowicz and Kroll (2007) showed that intermediate proficiency Spanish-English bilinguals were less accurate and slower in producing translation-ambiguous words compared to unambiguous words. The ambiguity disadvantage in reaction time was driven by abstract words, whereas translation of concrete words was less affected by the availability of multiple translations. Michael, Tokowicz, Degani, and Smith (2011) similarly found an accuracy disadvantage in producing translation-ambiguous words, and further showed that individual differences in working-memory and in the ability to ignore irrelevant information modulated these effects (see also Prior et al., 2013).

In a translation recognition task, in which participants are to determine whether a pair of words are correct translations of each other, Laxén and Lavaur (2010) showed that moderately proficient French-English bilinguals recognized translation ambiguous pairs less quickly and less accurately than translation unambiguous pairs. The effects were stronger when translation ambiguous words were presented with their less-dominant translation, and when the two possible translations were not related in meaning. A similar pattern emerged in a primed translation recognition task with moderately proficient English-German bilinguals (Eddington & Tokowicz, 2013).

Prior et al. (2013) demonstrated a similar effect of translation ambiguity with proficient Spanish/English bilinguals in both translation production and translation recognition tasks. Moreover, the cost associated with processing translation-ambiguous words could not be explained by lexical and semantic factors such as frequency, context availability, and cognate

status, because translation ambiguity impacted bilingual translation performance above and beyond these control variables. Further, the disadvantage observed for translation-ambiguous words was modulated by participants' proficiency, such that it decreased with increasing L2 proficiency.

Finally, even highly proficient bilinguals are influenced by translation-ambiguity. Boada et al. (2013) reported that highly proficient balanced Spanish/Catalan bilinguals were faster and more accurate to respond to translation-unambiguous pairs compared to translation-ambiguous pairs in a translation recognition task. The effects were observed for abstract and concrete words, and for cognate and non-cognate translations. Taken in concert, these results indicate that translation ambiguity is a crucial factor in understanding bilingual language processing.

1.3 The current study

In the current study we focus on the relation between within-language ambiguity (and in particular semantic ambiguity) and cross-language translation ambiguity. If within-language ambiguity underlies translation ambiguity, then the number of translations an English word receives in Dutch, for instance, should correlate with the number of translations that same word receives in Hebrew, Spanish, and German, because each source language meaning is likely to receive a different translation in every target language. Further, we measure whether the number of senses an English word has predicts the number of translations that same word receives in different target languages. In addition, for semantically ambiguous words, we examine if dominant meanings in the source language tend to be translated more often than subordinate meanings. If indeed source language semantic properties predict translation choices, this would provide evidence for semantic involvement in single-word translation processes (Brysbeart & Duyck, 2010; Guo, Misra, Tam, & Kroll, 2012; Kroll, van Hell, Tokowicz, & Green, 2010). In particular, assuming both direct lexical links between L2 and L1 words and conceptually mediated links between words in the two languages (e.g., The Revised Hierarchical Model, Kroll & Stewart, 1994), traces of source language semantic properties during translation would support reliance on conceptually mediated links. This is because translating an ambiguous word in the source language via the lexical route should activate the possible target language translations with equal likelihood, regardless of the meaning dominance in the source language. Conversely, translation relying on the conceptual route should be sensitive to meaning dominance in the source language, and would result in a higher probability of producing the translation that maps onto the more dominant meaning of the ambiguous source word.

2. Method

Four sets of previously collected number-of-translations norms were used, including English-to-Dutch (ED, Tokowicz et al., 2002), English-to-Spanish (ES, Prior et al., 2007), English-to-German (EG, Eddington et al., 2014), and English-to-Hebrew (EH, Smith et al., 2012) norms. All were collected with the same 'first-translation' method, in which each participant is requested to provide the first translation that comes to mind, and translation ambiguity is calculated by counting the number of different correct responses for each

source word provided across participants. Bilinguals of different language profiles participated in the different norming studies. EH participants were native Hebrew speakers who were advanced second language learners of English. Similarly, ED participants were native Dutch speakers, who were advanced second language learners of English. EG participants included native English speakers who were advanced second language learners of German, as well as native German speakers, who were advanced second language learners of English. ES participants included native English speakers who were advanced second language learners of Spanish, native Spanish speakers who were advanced second language learners of English, as well as several native speakers of both English and Spanish. See Table 1 for participant characteristics. The possible impact of this diversity in participant characteristics on the results is addressed in the discussion.

The four norming studies varied in the stimuli included. Of relevance, the ED (Tokowicz et al., 2002) and EG norms (Eddington et al., 2014) were collected for the same set of 561 English words. Similarly, the ES norms (Prior et al., 2007) and the EH norms (Smith et al., 2012) were collected for a single set of 670 English words. A subset of 208 English words was included in all four language pairs, allowing for direct comparisons in the magnitude of correlations in number of translations across all norms. Note, however, that the possible maximum number of translations varied somewhat across norms due to differences in the number of participants included in each of the norming studies (maximum of 6 in ED and EG, maximum of 10 in EH and maximum of 20 in ES). See Table 2 for stimulus characteristics

The focus of the current study was the correspondence between translation ambiguity and within-language semantic ambiguity. To this end, we used two measures of within-language ambiguity. First, we counted the number of senses each word received in WordNet (Fellbaum, 1998). This measure was available for all English words included in all four norming studies. We therefore calculated the correlation between number of English senses from WordNet and number of translations, for each set of norms independently.

In addition, we examined the correlation between within-language meaning dominance and translation dominance more closely. Because the different senses derived from WordNet tend at times to be highly related, it might overestimate the number of senses, and it is difficult to determine meaning dominance based on this measure. Thus, in order to identify the relative dominance of each meaning of items, we relied on an English association norms study (Twilley et al., 1994) that identified English words as ambiguous. A subset of these ambiguous words (n=126) were also included in two of the translation norms reported here (ES and EH), that relied on a large enough sample of participants to allow us to calculate the probability of the various translation choices (Prior et al., 2007). Translation probability was defined as the proportion of participants providing a specific translation out of all valid translations given to a particular source word, and in the current norms could range from 0.1 (for rarely produced translations) to 1 (for the single translation produced for unambiguous words). Such calculations were less suitable in the ED and EG norms due to the lower number of participants in those studies.

For each of the 126 homographs that appeared in both Twilley et al. (1994) association norms and in the Prior et al. (2007) and Smith et al. (2012) translation ambiguity norms, we matched the translation (in Spanish or in Hebrew) with the meaning it reflected in English. For instance, the translation 'rasha' in Hebrew was aligned with the homograph's meaning it reflected (i.e., the evil meaning of the homograph 'mean'). The alignment of a specific translation to a given meaning of an English word was determined by 2 proficient bilingual speakers of the languages in question (English-Spanish and English-Hebrew), and any discrepancies were discussed until agreement was achieved. We then predicted translation probability with within-language meaning probability (Twilley et al., 1994). Note that in these calculations, the number of items exceeds 126, because homographs that received more than one translation were included with each translation in the analysis. In the ES norms we were able to control for other factors that are known to affect translation probability (Prior et al., 2007) including target word length, log frequency, imageability, and the form similarity (cognate rating) of the target word and the source (English) word. Such control variables were not available for the EH norms. We further examined the correlations between translation probability in ES and translation probability in EH when the translations were aligned based on the English homograph meaning they reflected.

3. Results

The distributions of number of translations across the four sets of norms are visually illustrated in Figure 1. When interpreting these distributions it is important to keep in mind two key differences between the ED and EG norms on the one hand, and the ES and EH norms on the other hand. First, as mentioned earlier, the stimuli included in the ED and EG were initially identified as words having a single translation in previous psycholinguistic research (Tokowicz et al., 2002), whereas the stimuli of the ES and EH norms were not selected with this restriction. Second, the number of participants in the ED and EG norms was smaller than in the ES and EH norms, thus de facto limiting the variability and the maximum number of translations that could be provided for each English word. However, even taking these considerations into account, it is evident that there was a high proportion of translation ambiguity in all four sets of norms. As can be seen in Figure 1, at least 30% of items received more than one translation in all four norming studies.

As a second step, we examined to what degree the number of translations given for a specific English word in the four target languages was correlated. To this end, we analyzed the subset of 208 words that were included in all four norms. As can be seen in Table 3, there were significant moderate positive correlations between the number of translations provided for English words in Spanish, Hebrew, German, and Dutch. This suggests that when a given English word receives more than one translation in one target language, it also tends to receive more than one translation in other target languages. Because we examined target languages of various typological sources (e.g., Hebrew vs. German), such correlations are most likely not the result of similar characteristics in the target language (see also Tseng, Chang, & Tokowicz, 2014), but rather reflect a characteristic of the source English word, suggesting that ambiguity in the source language, be it semantic, morphological or related to part-of-speech, is a key driving force for translation ambiguity.

To specifically examine the contribution of within-language *semantic* ambiguity in the source language to translation ambiguity, we calculated the correlation between number of senses of the English words (from WordNet, Fellbaum, 1998) and the number of translations each word received in each of the four target languages. Supporting the role of within-language semantic ambiguity, the number of senses a word has in English significantly correlated with the number of translations it received in Dutch ($r = .098$, $N = 561$, $p = .021$), German ($r = .096$, $N = 561$, $p = .024$), Spanish ($r = .221$, $N = 670$, $p < .001$), and Hebrew ($r = .200$, $N = 670$, $p < .001$) (see Figure 2).[1] Note, however, that these correlations are not very strong, indicating that translation ambiguity cannot be reduced to within-language ambiguity alone, and likely reflects other sources such as near-synonymy in the target language.

To investigate further the contribution of within-language semantic ambiguity to translation ambiguity, we examined whether the probability of a specific meaning of an ambiguous English word (based on Twilley et al., 1994) influenced the probability of the Spanish and Hebrew translations corresponding to that meaning. Overall, we found that English meaning probability significantly and strongly correlated with translation probability in both Spanish ($r = .650$, $N = 172$, $p < .001$) and Hebrew ($r = .665$, $N = 173$, $p < .001$). Furthermore, translation probability in Spanish correlated with translation probability in Hebrew ($r = .715$, $N = 149$, $p < .001$), providing additional support to the strong role of source language meaning dominance in determining translation probability.

Finally, previous work has linked the probability of a given translation to its frequency and imageability, and to its form similarity with the word in the source language (Prior et al., 2007). Specifically, more frequent and imageable words are provided more often as translations than less frequent and imageable options. Similarly, cognate translations are given more often than their non-cognate counterparts (see also Prior et al., 2011). Here we examined whether within-language meaning probability, as determined by association norms to ambiguous words (Twilley et al., 1994), adds to the prediction of translation probability. This analysis was carried out only on the English to Spanish translation norms, because this was the only set of norms for which all predictor variables were available.

Using hierarchical regression, we entered meaning probability as the fourth step after controlling for target length and log frequency in the first step, image-ability in the second step, and form similarity (cognate overlap) in the third step (following Prior et al., 2007). Within-language meaning probability accounted for unique variance in translation probability ($\Delta R^2 = .036$, $N = 165$, $p = .01$) after controlling for previously-identified factors. Interestingly, as shown in Figure 3, meaning probability accounted for unique variance in translation probability both for English-dominant participants who translated from their first-language (L1) to their second (L2) ($\Delta R^2 = .029$, $N = 165$, $p = .023$) and for Spanish-

[1] An alternative way to examine the influence of source language semantic ambiguity on translation ambiguity is to examine the distribution of translations for ambiguous words with relatively unrelated senses (i.e., homographs, based on Twilley et al., 1994). Analysis of a subset of the items in each translation norming study that were identified as homographs in Twilley et al. (1994) reveal that indeed these ambiguous English words tend to receive more translations ($M_{ED} = 1.65$, $n_{ED} = 74$; $M_{EG} = 1.88$, $n_{EG} = 74$; $M_{EH} = 2.38$, $n_{EH} = 126$; $M_{ES} = 2.83$, $n_{ES} = 126$) than the remaining items in the norms ($M_{ED} = 1.37$, $n_{ED} = 487$; $M_{EG} = 1.53$, $n_{EG} = 487$; $M_{EH} = 1.86$, $n_{EH} = 544$; $M_{ES} = 1.98$, $n_{ES} = 544$).

dominant participants who translated from their L2 to their L1 (ΔR^2 = .036, N = 163, p = .01).[2] Direct comparisons between the two models are not possible, however, because each model predicted performance of a different population (i.e., different dependent variables). Nonetheless, it appears that proficient bilinguals are sensitive to the meaning probability of ambiguous words in both their L1 and in their L2 when performing a single-word translation task. It remains to be examined whether similar sensitivity is observed for less-proficient bilinguals, because prominent models of the bilingual lexicon predict developmental increases in access to meaning via L2 words (Kroll & Stewart, 1994). We return to this issue in the discussion.

4. General discussion

The current investigation demonstrates that patterns of translation ambiguity are consistent across several language pairs, including those that differ typologically. Specifically, the number of translations an English word received in Spanish, Dutch, German, and Hebrew correlated significantly. This highlights the role of ambiguity in the source language (English) as a driving force of translation ambiguity. To assess this relation more directly, we examined the correspondence between the number of senses of an English word (as a measure of within-language semantic ambiguity) and the number of translations that same word received in different target languages. The results show positive correlations, such that the more senses an English word has the more translations it tends to receive in other languages. A similar relation was recently observed with Japanese-English bilinguals, such that the more senses a word had in English, the more translations it tended to receive in Japanese (Allen & Conklin, 2014).

Thus, translation ambiguity is at least partially determined by within-language semantic ambiguity. Critically, however, within-language semantic ambiguity does not account for the full variability in translation ambiguity, because less than 5% of the variance in number of translations is explained by the number of senses the word has in the source language. This indicates that within-language semantic ambiguity is not the sole determiner of the existence of multiple translations. However, when the word is ambiguous in the source language, translation choices align tightly with meaning dominance. Specifically, the examination of the probabilities of specific translations in Spanish and Hebrew allowed us to characterize cross-language mapping in a more nuanced way. In particular, we demonstrated an alignment of the translation probabilities across the languages, driven by the semantic dominance structure of the English source word. Bilinguals more often translated the dominant meaning of an ambiguous English word than its subordinate meaning.

Recently, Arêas da Luz Fontes and Schwartz (2010) examined the relation between within-language meaning-dominance and cross-language form overlap. Specifically, they showed

[2]To alleviate concerns of interdependence between probabilities (e.g., when two meanings exist, the probability of one is fully determined by the probability of the other), which may artificially increase the correlations, we computed all of the above correlations with the highest probability meaning only (see also Prior et al., 2011). The results remained virtually the same, with meaning probability accounting for unique variance in both L1 to L2 (ΔR^2 = .033, N = 147, p = .021) and L2 to L1 translation (ΔR^2 = .039, N = 146, p = .011). Meaning probability significantly correlated with ES translation probability (r = 542, N = 122, p < .001) and with EH translation probability (r = .475, N = 121, p < .001). ES and EH translation probabilities correlated significantly (r = .635, N = 117, p < .001).

that when asked to produce sentences for English homographs, Spanish/English bilinguals tended to produce sentences to the meaning of the homograph that is also captured by the cognate translation in Spanish. For instance, because the word *arma* in Spanish captures only the 'weapon' meaning of the English homograph *arm* (and not its 'hand' meaning), bilinguals tended to generate sentences in English that fit the less dominant 'weapon' meaning of the word in English. These findings indicate that cross-language *form overlap* can bias within-language meaning dominance for bilinguals. In the current study we found that within-language meaning dominance predicted translation probability above and beyond form overlap and other lexical and semantic factors. Our results therefore extend the findings of Arêas da Luz Fontes and Schwartz (2010) by demonstrating that despite cognate-driven biases (see also Prior et al., 2013), the influence of form overlap does not override that of within-language meaning dominance.

Moreover, translation choice was affected by meaning dominance for target languages that differ dramatically from the source language. Specifically, translation choice in Hebrew was affected by English meaning dominance although the two languages differ typologically and do not share script. These findings converge with recent work of Allen and Conklin (2013) who show that performance of different script (Japanese-English) bilinguals is influenced by the number of senses the word has in the L2. Together, the relevant findings suggest an important role for within-language semantic ambiguity in determining translation ambiguity.

Participants in the four sets of norms included in the current study varied in their language profiles (see Table 1), with a wide range of L2 age of acquisition, proficiency, and immersion experience. The fact that we observed significant cross-norms correlations despite this variability attests to the robustness of the translation ambiguity effect. Indeed, the majority of the participants in the current study (all participants in the ED and EH norms, and half the participants in the EG and ES norms) provided translations in the L2 to L1 direction, and previous work suggests that L2 proficiency does not influence translation choice in this direction of translation (Prior et al., 2007). Because L2 proficiency does seem to influence L1 to L2 translation performance, such that lower proficiency in the L2 is associated with providing lower probability translations, it remains to be examined whether it similarly affects the link between within-language ambiguity and translation ambiguity in the forward direction of translation.

Interestingly, the influence of within-language semantic dominance structure on translation probability was similar in magnitude for English-dominant bilinguals translating into Spanish, their L2, and for Spanish dominant bilinguals translating from their L2 English into the L1. This finding indicates that moderately-to-highly proficient bilinguals are sensitive to the meaning structure of words in their L2, as expressed in their translation choices. According to models postulating that bilinguals translating from the L2 rely on lexical links to L1 translations (e.g., Jiang, 2000), translation choices are expected to be influenced by L1 lexical characteristics such as word frequency or the form overlap between the L1 and the L2. The finding that meaning dominance in the L2 influences translation choices above and beyond these L1 factors implies that these bilinguals access L2 meaning directly. This pattern is consistent with models postulating direct access to meaning from L2 words (e.g., Brysbaert & Duyck, 2010; for recent evidence see Guo et al., 2012), and a common

semantic/conceptual system subserving both languages (Kroll & Stewart, 1994; Van Hell & De Groot, 1998). Future research can track the development of meaning access through the L2 by examining the sensitivity of less proficient bilinguals to the meaning dominance structure when translating from L2 to L1.

To conclude, the current investigation sheds light on the sources of translation ambiguity, showing that it is partially, but not exclusively, driven by within-language semantic ambiguity. For meaning-ambiguous words, meaning dominance is a strong predictor of translation choice, going beyond the influence of form overlap and other lexical and semantic variables. The close alignment of meaning probability and translation probability supports the involvement of semantics when bilinguals process words in each of their languages.

Acknowledgments

During the writing of this manuscript, TD was supported by NSF- BCS-1052283, a Language Learning Dissertation Grant and the EU_FP7 grant CIG-322016. AP was supported by EU-FP7 grant IRG-249163. CME was supported by NSF DDIG 11-547 and the NIH MNTP R90DA023420. NT was supported by NIH-RO1-HD075800 and PSI2012-37623 awarded by the Spanish Ministry of Education and Science. The authors gratefully thank Prof. Ana I. Schwartz for her help with this project.

References

Allen D, Conklin K. Cross-linguistic similarity and task demands in Japanese-English bilingual processing. PloS One. 2013; 8:e72631. [PubMed: 24015266]

Allen D, Conklin K. Cross-linguistic similarity norms for Japanese-English translation equivalents. Behavior Research Methods. 2014; 46:540–563. [PubMed: 24046264]

Arêas Da Luz Fontes AB, Schwartz AI. On a different plane: Cross-language effects on the conceptual representations of within-language homonyms. Language and Cognitive Processes. 2010; 25:508–532.

Boada R, Sánchez-Casas R, Gavilán JM, García-Albea JE, Tokowicz N. Effect of multiple translations and cognate status on translation recognition performance of balanced bilinguals. Bilingualism: Language and Cognition. 2013; 16:183–197.

Brysbaert M, Duyck W. Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? Bilingualism: Language and Cognition. 2010; 13:359–371.

Brysbaert M, New B. Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. Behavior Research Methods. 2009; 41:977–990. [PubMed: 19897807]

Degani T, Tokowicz N. Ambiguous words are harder to learn. Bilingualism: Language and Cognition. 2010; 13:299–314.

Degani T, Tokowicz N. Cross-language influences: Translation status affects intra-word sense relatedness. Memory & Cognition. 2013; 41:1046–1064. [PubMed: 23658030]

Degani T, Tseng AM, Tokowicz N. Together or apart? Learning of translation ambiguous words. Bilingualism: Language and Cognition. 2014; 17:749–765.

Eddington CM, Tokowicz N. Examining English-German translation ambiguity using primed translation recognition. Bilingualism: Language and Cognition. 2013; 16:442–457.

Eddington CM, Degani T, Tokowicz N. English and German translation norms: Examining ambiguity and semantic variability of translations across proficiency levels. 2014 Manuscript in revision.

Fellbaum, C. WordNet: An Electronic Lexical Database. Bradford Books; 1998.

Francis, W. Bilingual semantic and conceptual representation. In: Kroll, JF.; DeGroot, AMB., editors. Handbook of Bilingualism,: Psycholinguistic approaches. New York: Oxford University Press; 2005. p. 251-267.

Frenck-Mestre C, Prince P. Second language autonomy. Journal of Memory and Language. 1997; 37:481–501.

Guo T, Misra M, Tam JW, Kroll JF. On the time course of accessing meaning in a second language: An electrophysiological investigation of translation recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2012; 38:1165–1186.

Jiang N. Lexical representation and development in a second language. Applied Linguistics. 2000; 21:47–77.

Kroll JF, Stewart E. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. Journal of Memory and Language. 1994; 33:149–174.

Kroll JF, Van Hell JG, Tokowicz N, Green DW. The Revised Hierarchical model: A critical review and assessment. Bilingualism: Language and Cognition. 2010; 13:373–381.

Laxén J, Lavaur JM. The role of semantics in translation recognition: Effects of number of translations, dominance of translations and semantic relatedness of multiple translations. Bilingualism: Language and Cognition. 2010; 13:157–183.

Marain V, Bartolotti J, Chabal S, Shook A. CLEARPOND: cross-linguistic easy access resource for phonological and orthographic neighborhood densities. PLoS One. 2012

Michael EB, Tokowicz N, Degani T, Smith CJ. Individual differences in the ability to resolve translation ambiguity across languages. Vigo International Journal of Applied Linguistics. 2011; 8:79–97.

Peterson RR, Savoy P. Lexical selection and phonological encoding during language production: evidence for cascaded processing. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1998; 24:539–557.

Prior A, Kroll JF, MacWhinney B. Translation ambiguity but not word class predicts translation performance. Bilingualism: Language and Cognition. 2013; 16:458–474.

Prior A, MacWhinney B, Kroll JF. Translation norms for English and Spanish: The role of lexical variables, word class, and L2 proficiency in negotiating translation ambiguity. Behavior Research Methods. 2007; 39:1029–1038. [PubMed: 18183923]

Prior A, Wintner S, MacWhinney B, Lavie A. Translation ambiguity in and out of context. Applied Psycholinguistics. 2011; 32:93–111.

Smith, Y.; Walters, J.; Prior, A. Translation norms for Hebrew and English. Edmond J. Safra Technical Report, 2012-1; 2012.

Tokowicz, N.; Degani, T. Translation ambiguity: Consequences for learning and processing. In: VanPatten, B.; Jegerski, J., editors. Research on second language processing and parsing. Amsterdam: John Benjamins; 2010.

Tokowicz N, Kroll JF. Number of meanings and concreteness: Consequences of ambiguity within and across languages. Language and Cognitive Processes. 2007; 22:727–779.

Tokowicz N, Kroll JF, de Groot AMB, van Hell JG. Number-of-translation norms for Dutch-English translation pairs: A new tool for examining language production. Behavior Research Methods, Instruments and Computers. 2002; 34:435–451.

Tseng, AM.; Chang, L-Y.; Tokowicz, N. Translation ambiguity between English and Mandarin Chinese: The roles of proficiency and word characteristics. In: Schwieter, J.; Ferreira, A., editors. The development of translation competence: Theories and methodologies from psycholinguistics and cognitive science. Cambridge: Cambridge Scholars Publishing; 2014. p. 107-165.

Twilley LC, Dixon P, Taylor D, Clark K. University of Alberta norms of relative meaning frequency for 566 homographs. Memory and Cognition. 1994; 22:111–126. [PubMed: 8035680]

Van Hell JG, De Groot AMB. Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. Bilingualism: Language and Cognition. 1998; 1:193–211.

Wilson MD. The MRC psycholinguistic database: Machine readable dictionary, version 2. Behavior Research Methods Instruments and Computers. 1988; 20:6–11.

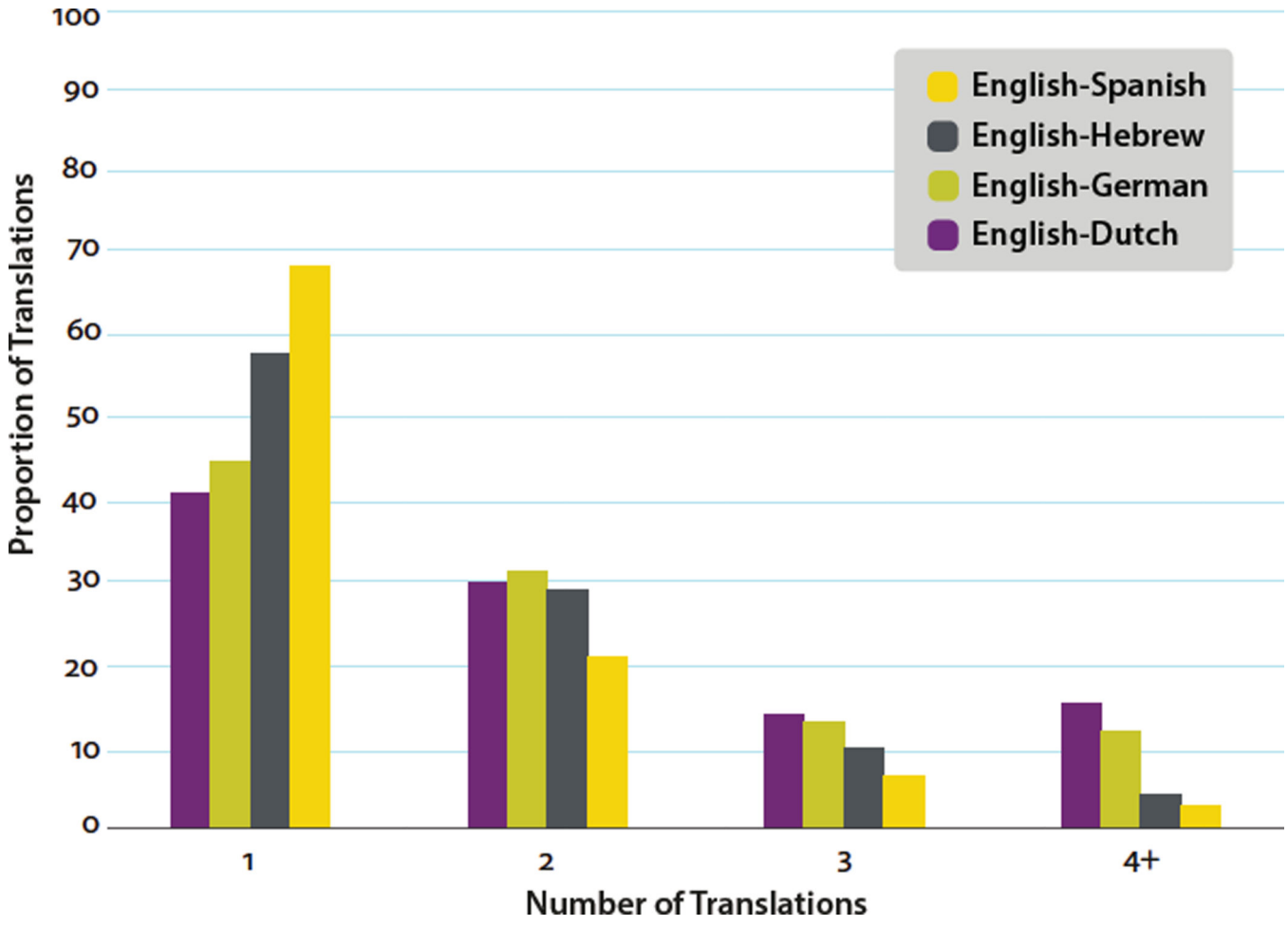

Figure 1.
Distribution of number of translations across the four norming studies.

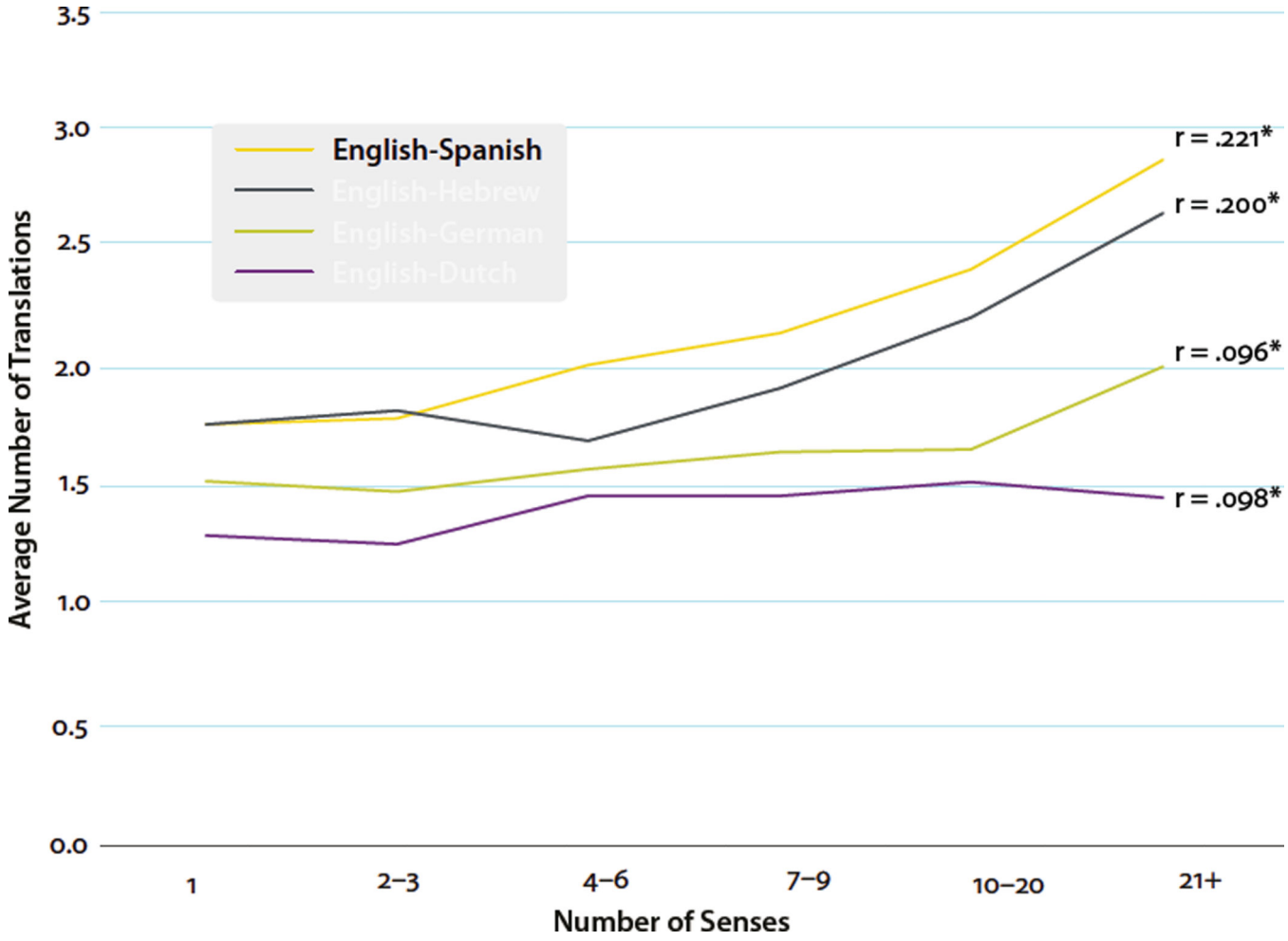

Figure 2.
Average number of translations by number of English senses across the four norming studies.

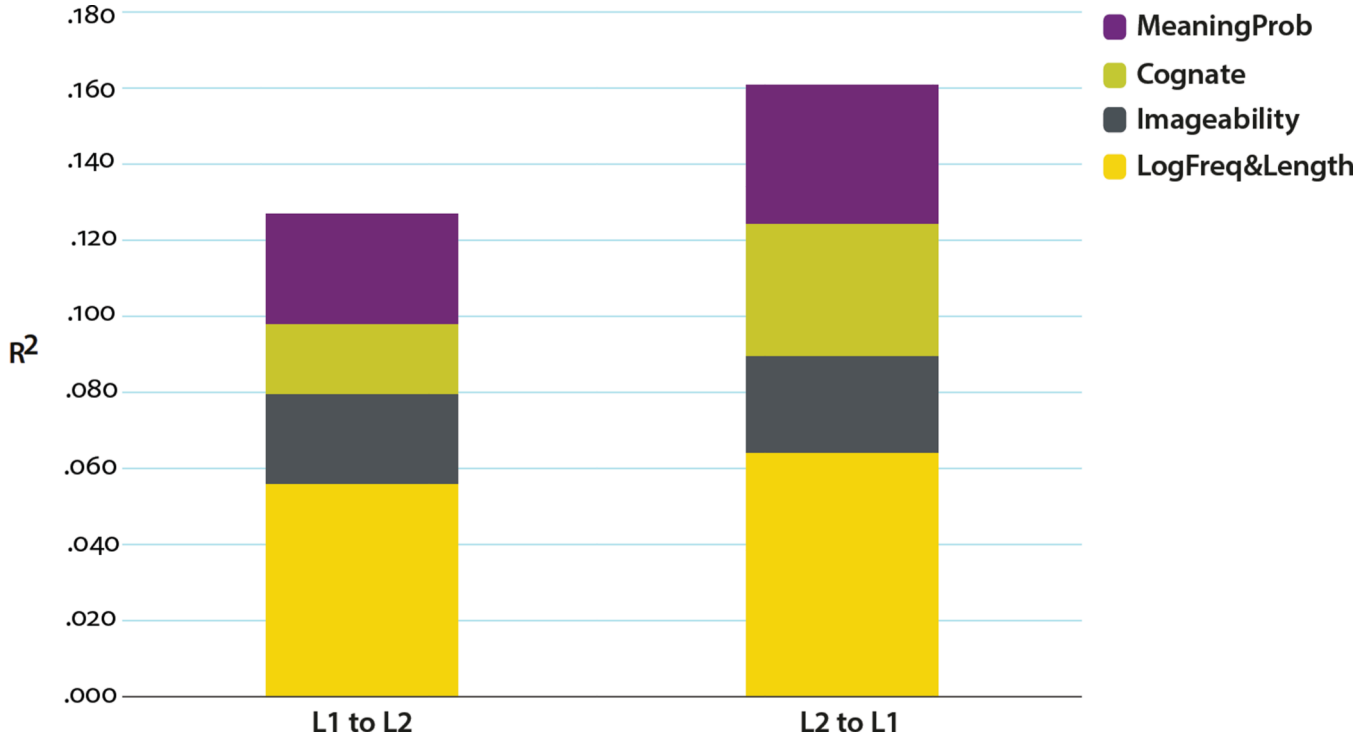

Figure 3.
Unique variance explained by within-language meaning probability in translation probability from English to Spanish, after entering control variables to the model.

Table 1

Participant characteristics based on self-reported Language History Questionnaires.

	Translation Norms			
	English-Dutch	**English-German**	**English-Spanish**	**English-Hebrew**
Total N [n per word]	12 [6]	6 [6]	40 [20]	42 [10]
Age (years)	21.6 (3)	25.1 (2.2)	30.9 (8.4)	29.5 (6.8)
L2	English	English (n=3) / German (n=3)	English (n=10) / Spanish (n=10) *	English
L2 Age of Acquisition	10.5 (2.1)	14.8 (3.6)	12.7 (6.9)	7.7 (3.4)
L2 immersion experience in months	5 (12.3)	8 (5.2)	32.2 (35.7)	13.9 (23.8)
L2 proficiency, 1–10 scale	7.2 (1.2)	7.5 (1.3)	7.9 (1.3)	8.1 (1.1)
English proficiency, 1–10 scale	7.2 (1.2)	8.8 (0.8)	9.0 (1.2)	8.1 (1.1)

Note. [n per word] refers to the number of different participants who provided translations for each item because each participant translated a portion of the stimulus list. Proficiency scores are the average of self-rated proficiency in reading, writing, speaking and comprehension on a 1–10 scale, where 1 indicates the lowest level of proficiency and 10 indicates the highest level of proficiency. Standard deviations are shown in parentheses.

* L2 (English vs. Spanish) was determined based on dominance rather than age of acquisition in the ES norms.

Table 2

Stimulus characteristics.

	Translation Norms							
	English-Dutch & English-German		**English-Spanish & English-Hebrew**		**Shared Set Across Norms**			
Number of Items	561		670		208			
Length (in letters)	5.3 (1.8)		5.4 (1.9)		5.0 (1.6)			
Subtlex Frequency per million	76.1 (176.4)		202.8 (614.8)		118.2 (194.5)			
Concreteness Ratings (100–700)	495.0 (122.9)		457.4 (124.0)		504.7 (120.8)			
Number of Senses (WordNet)	6.1 (4.8)		9.1 (8.8)		7.4 (5.2)			
Number of Translations in Target Language	ED	EG	ES	EH	ED	EG	ES	EH
	1.4 (.7)	1.6 (.8)	1.4 (1.2)	2.0 (1.2)	1.3 (.7)	1.4 (.6)	1.9 (1.2)	1.6 (.9)

Note Frequency counts are based on the Subtlex corpus (Brysbaert & New, 2009) and were extracted through Clearpond (Marain, Bartolotti, Chabal, & Shook, 2012). Concreteness ratings were taken from the MRC database (Wilson, 1988). The number of English senses was taken as the number of definitions in WordNet (Fellbaum, 1998). Standard deviations are shown in parentheses.

Table 3

Correlations between the number of translations given to English words in Spanish (ES), Hebrew (EH), German (EG), and Dutch (ED).

Number of Translations in	ES	EH	EG	ED
ES	1	–	–	–
EH	.329*	1	–	–
EG	.283*	.334*	1	–
ED	.391*	.305*	.312*	1

*Correlation is significant at the 0.01 level (2-tailed). Listwise N=208