Running head: DUTCH-ENGLISH NORMS

Number of Translation Norms for Dutch-English Translation Pairs: A New Tool for Examining Language Production

Natasha Tokowicz

Carnegie Mellon University

Judith F. Kroll

The Pennsylvania State University

Annette M. B. de Groot

University of Amsterdam

and

Janet G. van Hell

University of Nijmegen

Address for Correspondence:

Natasha Tokowicz
Department of Psychology
342C Baker Hall
Carnegie Mellon University
Pittsburgh, PA 15213
Phone: 412-268-6745
Fax: 412-268-8215
E-mail: Tokowicz@cmu.edu

Abstract

We collected number of translations norms on 562 Dutch-English translation pairs from several previous studies of cross-language processing. Participants were highly proficient Dutch-English bilinguals. Form and semantic similarity ratings were collected on the 1003 possible translation pairs for the original word pairs that were generated by the norming participants. At least 40% of the words in each direction of translation had more than one translation. Approximately 40% of the translations were rated as being dissimilar with respect to spelling/sound across languages (i.e., were 'noncognates'). Approximately 45% of the translations were rated as being highly semantically similar across languages. The form similarity ratings were found to be highly reliable even when obtained using different bilinguals and modified rating procedures. Number of translations and meaning factors significantly predict semantic similarity of translation pairs. In future research, these norms may be used to determine the number of translations of words to control for or study this factor.

Number of Translation Norms for Dutch-English Translation Pairs: A New Tool for Examining Language Production

Experimental research on language production has relied extensively on picture naming paradigms to reveal the processes that lead up to the articulation of a spoken word (e.g., Jescheniak & Schriefers, 1998; Levelt, Roelofs, & Meyer, 1999; Peterson & Savoy, 1998; Starreveld & La Heij, 1995). The development of norms for picture materials across a variety of languages allows for control of the language-specific properties of pictures' names, such as name agreement (the degree to which a picture is labeled with the same name, e.g., Alario & Ferrand, 1999, for French; Sanfeliu & Fernandez, 1996, for Spanish; Snodgrass & Vanderwart, 1980, for English). These measures allow for more direct comparison across experiments performed with different language speakers and also for control of the properties of spoken responses in production experiments with bilingual speakers. However, a goal of research on language production is to understand how spoken utterances are generated for a wide range of concepts, not only those that can be depicted as simple line drawings.

In the within-language domain, a number of alternative tasks have been devised to achieve a more general characterization of lexical access in production, including a definition naming task first reported by La Heij, Starreveld, & Steehouwer (1993). In the cross-language domain,

translation production, a task in which a word is presented in one language and its translation equivalent must be spoken in the other language, is frequently used to examine language production. Potter, So, Von Eckardt, & Feldman (1984) demonstrated that proficient bilinguals were as fast to translate from their first language (L1) to their second (L2) as to name pictures in L2. La Heij et al. (1990) then showed that the translation task could also be modified into a Stroop-type task, similar to the picture-word interference task, by adding distractor words that were related to the meaning or to the form of the word to be spoken. The findings in these initial experiments on translation and picture naming suggested that the two tasks produced the same pattern of results and therefore could be used relatively interchangeably. However, because the words in these translation experiments were the names of the objects to which they referred in the picture naming tasks, they were subject to the same limitations based on the modest number of easily pictured nouns.

More recently, a series of studies has examined the processes underlying translation production without direct comparison to picture naming so that the words in the translation task reflect a more representative distribution of spoken language (De Groot, 1992; De Groot, Dannenburg, & Van Hell, 1994; Kroll, Michael, Tokowicz, & Dufour, in press; Sánchez-Casas, Davis, & García-Albea, 1992; Tokowicz, 2000; Van Hell, 1998). However, the name agreement problem that arises in picture

naming is also a problem for translation; some words correspond to a number of alternative translations in the other language, whereas other words have only a single dominant translation. The number of translations that words have across languages is a factor that has recently been shown to affect translation performance (e.g., Kroll & Tokowicz, 2001; Schönpflug, 1997; Tokowicz, 2000; Tokowicz & Kroll, 2001). Indeed, the effect of the number of translations variable is often reported as being larger than that of other variables that also influence translation performance. Importantly, number of translations is potentially related to many within-language factors. Therefore, confounding of variables could occur if normative measures are not used to determine the number of translations of stimuli. The goal of the present paper is to provide such a set of measures for a large number of Dutch-English translation pairs that have been used in several different studies. These norms will allow researchers to control for number of translations or to study the impact of number of translations on cross-language processing. Furthermore, these norms could allow words with multiple translations to be utilized in a manner similar to pictures with multiple names (e.g., Peterson & Savoy, 1998) to study language production in bilinguals.

Number of Translations

Many bilinguals report that not all words in each language can be readily translated into the other language. Indeed, this observation has

been reported as evidence for linguistic relativity (e.g., Green, 1998; Pavlenko, 1999). Some words fail to have a direct translation equivalent. For example, in Dutch, the word "gezelligheid" means a warm and cozy feeling that doesn't map directly to any of its English translations, such as coziness, hominess, snugness, or domesticity. In contrast, the Dutch word "appel" is "apple" in English and the correspondence is direct and bidirectional. However, many words in one language have more than a single appropriate translation in the other language.

Multiple translations across languages can arise in a number of different ways. First, lexical ambiguity within a language (e.g., due to synonymy) can lead to the availability of multiple translations for many words across languages. For example, the Dutch word "herfst" translates to either "autumn" or "fall" in English. Furthermore, some words have more than a single meaning within a language. For example, the English word "trunk" has several possible translations in Dutch depending on the meaning of the word that is being translated (the car meaning translates to "achterbak" or "kofferbak", the tree meaning translates to "stam" or "boomstronk", the body meaning translates to "romp", and the snout meaning translates to "slurf"). Thus, perhaps any one would be considered a less appropriate translation of the word (overall) because it does not encompass all meanings of the word.

To examine this issue, we first obtained number of translations

norms on the 440 Dutch-English translation pairs that have been used by De Groot and her colleagues in several past studies (e.g., De Groot, 1992; De Groot, Borgwaldt, Bos, & Van Den Eijnden, in press; De Groot et al., 1994; Van Hell, 1998)[1] as well as an additional 122 English words used by Dijkstra, Grainger, and Van Heuven. (1999) and their Dutch translations. We then obtained semantic similarity (i.e., semantic differential, Lambert, Havelka, & Crosby, 1958) norms on the entire set of 1003 possible translation pairs that were generated by the norming participants for the original set of Dutch and English words (i.e., each word paired with each of its possible translations). Thus, it was possible to examine the relation between number of translations and semantic similarity across languages. Furthermore, because these materials included translation pairs that varied in their concreteness (e.g., "desk" is more concrete than "idea"), it was possible to examine the hypothesis (De Groot et al., 1994) that concrete word translations are more semantically similar across languages than abstract word translations.

Form Similarity

Although the number of translations of a word may affect the semantic similarity of that word to its translation(s), there are other factors that may influence the semantic similarity of translation pairs. Another factor hypothesized to influence semantic similarity of translation equivalents is form similarity across languages; De Groot (1992 ; De Groot

et al., 1994) proposed that translation pairs higher in their form similarity are also more similar in meaning across languages. We therefore obtained form (spelling/sound) similarity ratings on all possible translation pairs. Although form similarity ratings were available for the 440 word pairs on which norms had been obtained by De Groot et al. (in press), our number of translations norming study showed that other possible translations were given for some of the words in each direction of translation. Therefore, it was necessary to obtain form similarity norms on all possible translation pairs.

In addition to its potential influence on semantic similarity of translation equivalents, form overlap has been shown to affect bilingual language performance on a variety of tasks including translation, word naming, picture naming, and lexical decision. In general, the more similar translations are across languages, the more quickly and accurately they are processed. However, relatively few sources for normative data on this measure are readily available. As a result, many researchers use one of three alternative means for distinguishing cognate from noncognate translations. First, some researchers only classify translations that have identical orthography and similar phonology across languages as cognates (e.g., bed-bed in English and Dutch; e.g., Gerard & Scarborough, 1989). This method typically results in striking effects of cognate status on the measure of interest, but ignores the relatively large group of words that

are similar in orthography and phonology across languages (e.g., apple-appel in English and Dutch), as well as words that have regular orthographic changes from one language to the other (e.g., -ty in English regularly changes to –teit or –heid in Dutch, as in university-universiteit). Another common method for classifying cognates is to have participants view words in a language with which they are not familiar and give a possible translation; then, words that are typically assigned the correct translation at least a certain amount of the time are classified as cognates (e.g., Kroll & Stewart, 1994). However, naïve participants may not be aware of the regular changes between languages and also are unlikely to be familiar with the phonology of the other language; this knowledge is likely to influence bilingual language processing (see Friel & Kennison, in press, for a comparison of several cognate identification techniques, including cognate norms collected from naïve participants who were provided with other-language phonology information). The final approach that has been taken is for the experimenter to assign cognate status to translation pairs. This approach is unfavorable because researchers have unique knowledge of existing theories of language representation and processing. A further criticism of all of the aforementioned approaches is that they ignore the continuous nature of similarity that could be used to predict performance on cross-language tasks. Here, we report continuous norms on a large set of translations that

can be used by researchers in future studies.

Method

Word Sample

The original sample of 562 word pairs used in this normative study included two different sets of items that had been used in several past experiments of bilingual language processing (De Groot, 1992; De Groot et al., 1994; De Groot et al., in press; Dijkstra et al., 1999; Van Hell, 1998; Van Hell & De Groot, 1998a, 1998b). Subsequently, all possible translation pairs generated from the number of translations norming study were rated on cognate and semantic similarity (i.e., all words paired with all other possible translations). These norms are available at http://www.talkbank.org/natasha/norms/.[2]

Number of Translations Norms

Number of translations norms were obtained to determine the absolute number of distinct translations that were assigned to each word presented in isolation (e.g., an English word presented without a Dutch translation). From these translations it was possible to determine the number of distinct meanings of a word that had been translated and the number of translations (or lexical forms) associated with each of those meanings. Note that although we derived a measure we refer to as 'number of meanings' from our norming study, this number of meanings is only suggestive of the absolute number of meanings of the word within

the language, and rather should be considered the number of meanings that have consequences for translation across languages.

Participants. The participants in the number of translations norming task were 24 Dutch-English bilinguals from the population of Dutch university students at the University of Nijmegen, Nijmegen, The Netherlands. The participants completed language history questionnaires in which they reported on their second language learning experiences. The participants rated their first and second language reading, writing, conversational, and speech comprehension ability on a scale that ranged from one to ten, and indicated the age at which L2 learning began and the types of exposure they had learning the L2; see Table 1 for the language history questionnaire data.

Procedure. The words were divided into several list versions and printed in random order in booklets. Each participant wrote their first spontaneous translation for each word and translated words from only one version and into only one language.

Scoring. The responses were coded for accuracy using Prisma Dutch-English and English-Dutch dictionaries (1990) and by a native Dutch speaker who was an advanced student of English at the University of Nijmegen. The number of correct possible translations for each word was calculated based on the total number of expected translations (those that had been considered correct in the original experiment from which

the stimuli were drawn), synonyms, other meanings (when a different meaning of the stimulus word was translated), verb meanings, and colloquial uses. The number of possible translations for each word in each direction of translation and the number of meanings to which the translations corresponded for a given word in a given direction were calculated.

Form Similarity and Semantic Similarity Norms

Ratings of form similarity and semantic similarity were obtained for all correct translation pairs that had been given in the number of translations norming task (i.e., not only for the original translation pair, but also for each stimulus paired with every other possible translation).

Participants. These ratings were obtained from 16 Dutch-English bilinguals from the same population as those in the number of translations norming experiment (see Table 1 for their language history questionnaire data). Participants in the similarity rating study had not participated in the number of translations norming study.

Procedure. Participants rated the similarity of the words in each translation pair in terms of (1) their meaning similarity, and (2) their combined spelling and sound similarity; both ratings were performed on a 7 point Likert-type scale where 1 indicated low similarity and 7 indicated high similarity (e.g., De Groot & Nas, 1991; see the Appendix for the rating instructions). The instructions indicated that the participants were

first to complete the semantic similarity rating for a pair and then complete the spelling/sound similarity rating. The semantic similarity rating was done prior to the form similarity rating because pilot participants on the form similarity rating task reported difficulty rating the word pairs with respect to their spelling and sound similarity without taking their meaning similarity into account.

Results and Discussion

Number of Translations and Number of Meanings Translated

The number of translations data show that most of the translations given were expected responses (see Table 2). Most of the words were assigned only one translation, suggesting that the stimulus set had a relatively large proportion of unambiguous stimuli (see Figure 1). In both directions, the range was from zero (i.e., an acceptable translation was never given) to five. At least 20 per cent of words in each direction of translation were assigned two translations, which makes these norms useful for selecting stimuli that can be used to examine the effects of multiple translations. These data also demonstrate that significantly more translations were given from English to Dutch than Dutch to English, ($M_{\text{Dutch to English}} = 1.29$, $SD = .58$; $M_{\text{English to Dutch}} = 1.40$, $SD = .77$), $t\,(561) = 3.1$, $p < .01$, which most likely reflects the participants' larger vocabulary in Dutch.

The number of meanings that were translated from Dutch to

English and English to Dutch are shown in Figure 2. In both directions, the range was from zero to three, and most of the words had only one meaning translated. However, almost ten per cent of the words in each direction had multiple meanings translated. Therefore, these norms can be used to select stimuli that vary on this dimension so that the effects of number of meanings can be measured.

Form Similarity and Semantic Similarity Ratings

The rating results reflect the data from the set of all 1003 possible translation pairs that were generated from the original set of 562 translation pairs. The form similarity ratings encompassed the full range of the rating scale and the mean rating was at the dissimilar end of the scale ($M = 2.6$, $SD = 1.9$), thus, there were many items that could be considered noncognates. Furthermore, relatively few items received similarity ratings that were within the high similarity half of the range (see Figure 3 for the distribution of the mean ratings).

Although there was a statistically significant difference between the mean ratings from English to Dutch and Dutch to English overall, t (1002) $= 7.8$, $p < .01$, the correlation between them was very high (the English to Dutch ratings were slightly higher than the Dutch to English ratings; $M_{\text{English to Dutch}} = 2.7$, $SD = 2.0$, $M_{\text{Dutch to English}} = 2.6$, $SD = 1.9$), $r = .95$, $p < .01$. Therefore, the average across the directions will be used in subsequent analyses.

The mean semantic similarity ratings were at the similar end of the scale, which is to be expected because all of the word pairs were translation equivalents (M = 6.4, SD = .7; see Figure 4 for the distribution of the mean semantic similarity ratings). Furthermore, no pair received a mean rating lower than 2.5. As for the form similarity ratings, there was a statistically significant difference between the mean ratings in the two directions, such that the English to Dutch ratings were slightly higher than the Dutch to English ratings ($M_{English\ to\ Dutch}$ = 6.44, SD = .76, $M_{Dutch\ to\ English}$ = 6.36, SD = .83), t (1002) = 3.7, p < .01. The correlation between the ratings in the two directions was significant, r = .66, p < .01. Therefore, the average across the directions will be used in subsequent analyses.

Reliability of Form Similarity Ratings

De Groot (1992) obtained norms on a set of Dutch-English translation pairs for form (combined spelling and sound) overlap across languages from a group of students at the University of Amsterdam, Amsterdam, The Netherlands. De Groot reported that the words for which norms were obtained in both studies showed a high correlation (r = .98). We obtained form similarity norms for a large set of words, a subset of which were the 440 translation pairs included in the De Groot study. We also found the correlation between the two sets of ratings to be highly significant, r = .98, p < .01. This indicates that although the ratings were collected from different (but similar) populations of bilinguals at different

times (approximately ten years apart), they are highly reliable. Furthermore, the high correlation suggests that performing the semantic similarity rating prior to the form similarity rating did not change the form similarity ratings significantly.

Correlations between Form Similarity, Number of Translations, Concreteness, Context Availability, and Semantic Similarity

We examined the intercorrelations among the factors for which norms on the set of 440 translation pairs were obtained or already available (De Groot et al., in press). The factors included in this correlational analysis are form similarity, number of translations, concreteness, context availability, and semantic similarity. These correlations were performed using the mean form and semantic similarity ratings across languages (see Table 3).

Of particular interest is the correlation between semantic similarity and several other measures. It was hypothesized that words with more than one translation may be considered less similar to any one of their translations in meaning than words with only a single translation. Indeed, semantic similarity is significantly negatively correlated with the number of distinct translations a word has, such that the higher the number of translations, the lower the similarity of the translation pair with respect to meaning. In fact, studies that have examined the effects of number of translations on translation performance have shown that words with

multiple translations are translated more slowly and less accurately than words with only one translation (Kroll & Tokowicz, 2001; Schönpflug, 1997; Tokowicz, 2000; Tokowicz & Kroll, 2001). One possible explanation for this finding is that the availability of multiple translations increases competition prior to the selection of a single lexical candidate for output. Another possible explanation for slower and less accurate responses for multiple translation words is that they are less semantically similar than single translation words.

De Groot (1992; De Groot et al., 1994) has hypothesized that words higher in concreteness are more semantically similar to their translations across languages than words low in concreteness. The data are consistent with this suggestion; semantic similarity is correlated with word concreteness, such that the higher the concreteness of the word, the higher the meaning similarity of that word and its translation. Context availability reflects the ease with which context can be accessed for a word, and is typically highly correlated with concreteness, such that words higher in concreteness are also higher in their context availability (rs = .73 and .65, Schwanenflugel, Harnishfeger, & Stowe, 1988, Experiments 2 & 3, respectively; r = .88, Schwanenflugel & Shoben, 1983). In the De Groot et al. (in press) norms, concreteness and context availability are also significantly correlated. Furthermore, context availability is correlated with semantic similarity such that the higher the

context availability, the more semantically similar the word is to its translation.

Semantic similarity does not correlate significantly with form similarity across languages. This finding is not consistent with De Groot's (1992) hypothesis that words which share form across languages are also more likely to share meaning across languages.

Although these correlations are suggestive of the relations among the factors of interest in this study, they are ambiguous as to which factors are responsible for predicting unique variance in semantic similarity ratings. Therefore, we used a hierarchical linear regression analysis to partition the variance in semantic similarity that is uniquely attributable to concreteness/context availability, form similarity, and number of translations. More specifically, we used this technique to determine whether the relatively new variable, number of translations, has any unique predictive power above and beyond that of concreteness/context availability and form similarity, which have been studied more extensively in past research.

We entered the concreteness and context availability ratings (in Dutch and English) on the first step of the analysis, the form similarity ratings on the second step of the analysis, and the number of translations (from Dutch to English and English to Dutch) on the final step of the analysis; these factors were used to predict semantic similarity rating.

The first step showed that concreteness and context availability accounted for a moderate proportion of the variance in semantic similarity, such that higher concreteness/context availability was associated with higher semantic similarity, $R^2 = .14$, F (4, 433) = 17.7, $p < .01$. Within this first step, only Dutch and English context availability were significant contributing factors, $t_{Dutch} = 2.6$, $p < .05$, and $t_{English} = 4.4$, $p < .01$.

After controlling for concreteness/context availability in the first step of the analysis, form similarity did not account for an additional proportion of the variance in semantic similarity in the second step, $\Delta R^2 = .00$. However, after accounting for concreteness/context availability and form similarity in the first two steps of the analysis, number of translations accounted for a significant proportion of the remaining variance in semantic similarity, $\Delta R^2 = .04$, $F\Delta$ (2, 430) = 10.9, $p < .01$. Within this third step, the number of translations from English to Dutch accounted for a significant proportion of unique variance, such that the greater the number of translations, the lower the semantic similarity, $t = 4.4$, $p < .01$; Dutch and English context availability still accounted for significant proportions of variance. Thus, the results from the hierarchical regression analysis were consistent with the conclusions drawn on the basis of the previous correlations. Specifically, context availability and number of translations account for unique proportions of variance in semantic similarity ratings, whereas concreteness and form similarity do

not.

Conclusions

Like the information available regarding name agreement for pictured objects, the present set of norms make it possible for researchers using Dutch-English translation pairs to select stimuli on the basis of number of translations and/or similarity. Because the majority of bilingual translation studies are conducted on Dutch-English bilinguals who are a relatively homogeneous group of highly proficient bilinguals, we believe these norms will be useful in many future studies. We obtained number of translations norms on a set of 562 Dutch-English translation pairs. We then obtained form and semantic similarity ratings on the set of 1003 possible Dutch-English translation pairs that were generated during the number of translations norming task. Overall, the majority of the words had only a single translation across languages from both Dutch to English and English to Dutch. Furthermore, most of the translation pairs were rated as being relatively dissimilar with respect to form and relatively similar with respect to meaning. Correlations showed that the form similarity ratings we collected are highly correlated with those obtained in past norming studies. Intercorrelations among the factors showed that semantic similarity is correlated with concreteness and context availability, such that the higher the concreteness or the more available the context, the more similar the translation pairs are in

meaning. Also, the more translations a word has, the lower the semantic similarity of the translation pair. Finally, semantic similarity is not correlated with form similarity across languages. These correlational results were confirmed using hierarchical regression analyses. These results may benefit future research because we report normative data on a measure that has not been studied extensively in past research (number of translations), but has been shown to affect the semantic similarity of translation pairs as well as translation performance.

References

Alario, F. X., & Ferrand, L. (1999). A set of pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. Behavior Research Methods, Instruments, and Computers, 31, 551-552.

De Groot, A. M. B. (1992). Determinants of word translation. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 1001-1018.

De Groot, A. M. B., Borgwaldt, S., Bos, M., & van den Eijnden, E. (in press). Lexical decision and word naming in bilinguals: Language effects and task effects. Journal of Memory and Language.

De Groot, A. M. B., Dannenburg, L., & Van Hell, J. G (1994). Forward and backward word translation by bilinguals. Journal of Memory and Language, 33, 600-629.

De Groot, A. M. B., & Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. Journal of Memory and Language, 30, 90-123.

Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. Journal of Memory and Language, 41, 496-518.

Friel, B. M., & Kennison, S. M. (in press). Identifying German-English cognates, false cognates, and non-cognates: Methodological issues

and descriptive norms. Bilingualism: Language and Cognition.

Gerard, L. D., & Scarborough, D. L. (1989). Language-specific lexical access of homographs by bilinguals. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 305-313.

Green, D. W. (1998). Bilingualism and thought. Psychologica Belgica, 38, 251-276.

Jescheniak, J. D., & Schriefers, H. (1998). Discrete serial versus cascaded processing in lexical access in speech production: Evidence from the coactivation of near-synonyms. Journal of Experimental Psychology: Learning, Memory, and Cognition, 24, 1256-1274.

Kroll, J. F., Michael, E. B., Tokowicz, N., & Dufour, R. (in press). The development of lexical fluency in a second language. Second Language Research.

Kroll, J. F., & Stewart, E. (1994). Concept interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. Journal of Memory and Language, 33, 149-174.

Kroll, J. F., & Tokowicz, N. (2001). The development of conceptual representation for words in a second language. In J. L. Nicol (Ed.), One mind, two languages: Bilingual language processing (pp. 49-71). Malden, MA: Blackwell Publishers.

La Heij, W., De Bruyn, E., Elens, E., Hartsuiker, R., Helaha, D., &

Van Schelven, L. (1990). Orthographic facilitation and categorical interference in a word-translation variant of the Stroop task. Canadian Journal of Psychology, 44, 76-83.

La Heij, W., Starreveld, P. & Steehouwer, L. (1993). Semantic interference and orthographic facilitation in definition naming. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 352-368.

Lambert, W. E., Havelka, J., & Crosby, C. (1958). The influence of language acquisition contexts on bilingualism. Journal of Abnormal Social Psychology, 56, 239-244.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999) A theory of lexical access in speech production. Behavioral and Brain Sciences, 22, 1-75.

Pavlenko, A. (1999). New approaches to concepts in bilingual memory. Bilingualism: Language and Cognition, 2, 209-230.

Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. Journal of Experimental Psychology: Learning, Memory, and Cognition, 24, 539-557.

Potter, M. C., So, K-F., Von Eckhardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and more proficient bilinguals. Journal of Verbal Learning and Verbal Behavior, 23, 23-38.

Prisma Woordenboek: Engels-Nederlands (1990). Utrecht, The

Netherlands: Uitgeverij Het Spectrum.

Prisma Woordenboek: Nederlands-Engels (1990). Utrecht, The Netherlands: Uitgeverij Het Spectrum.

Sánchez-Casas, R. M., Davis, C. W., & García-Albea, J. E. (1992). Bilingual lexical processing: Exploring the cognate/non-cognate distinction. European Journal of Cognitive Psychology, 4, 293-310.

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.

Schönpflug, U. (1997, April). Bilingualism and memory. Paper presented at the first International Symposium on Bilingualism, Newcastle-upon-Tyne, United Kingdom.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. Journal of Memory and Language, 27, 499-520.

Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 82-102.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity,

and visual complexity. Journal of Experimental Psychology: Human Learning & Memory, 6, 174-215.

Starreveld, P. A., & La Heij, W. (1995) Semantic interference, orthographic facilitation, and their interaction in naming tasks. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21, 686-698.

Tokowicz, N. (2000). Meaning representation within and across languages. Unpublished dissertation, The Pennsylvania State University.

Tokowicz, N., & Kroll, J. F. (2001). Accessing meaning for words in two languages: The effects of concreteness and multiple translations in bilingual production. Manuscript in preparation.

Van Hell, J. G. (1998). Cross-language processing and bilingual memory organization. Unpublished dissertation, University of Amsterdam.

Van Hell, J. G. , & De Groot, A. M. B. (1998a). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. Bilingualism: Language and Cognition, 1, 193-211.

Van Hell, J. G., & De Groot, A. M. B. (1998b). Disentangling context availability and concreteness in lexical decision and word translation. Quarterly Journal of Experimental Psychology (A), 51, 41-63.

Appendix

Instructions for Cognate Ratings (Adapted from De Groot & Nas, 1991)

Many words in Dutch and English share sound and/or spelling in the two languages. The following word pairs consist of a Dutch word and one of its English translations.

Your first task is to rate the similarity of the two words in terms of their meaning. The rating scale goes from 1, which indicates "completely different" to 7, which indicates "exactly the same".

Your second task is to rate the similarity of each word pair in terms of spelling and sound. The rating scale goes from 1, which indicates "low similarity" to 7, which indicates "high similarity". <u>Your rating should reflect a combination of both the spelling and sound similarity.</u>

Examples:

		Meaning							Spelling/Sound						
		completely different				exactly the same			low similarity				high similarity		
jurk	dress	1	2	3	4	5	6	**7**	**1**	2	3	4	5	6	7
pen	pen	1	2	3	4	5	6	**7**	1	2	3	4	5	6	**7**
arrival	citroen	**1**	2	3	4	5	6	7	**1**	2	3	4	5	6	7

If you are not sure how to rate a word pair, it is appropriate to guess or follow your first instinct. Please rate the items in the order in which they appear in the list. Rate each pair in terms of meaning first, then spelling/sound. Please do not change your responses or go back to a previous item.

Author Notes

This research was supported in part by a teaching release grant to Natasha Tokowicz from the Research and Graduate Studies Office at The Pennsylvania State University, by NSF Grants BCS-9905850 and BCS-0111734, and by a sabbatical grant from the Nijmegen Institute for Cognition and Information, The Netherlands, to Judith F. Kroll. The norming data were collected as part of the Ph.D. thesis awarded to the first author from The Pennsylvania State University. We thank Roel Kerkhofs for collecting the form and semantic similarity norms and Marc Rensen for his invaluable assistance with data coding of the number of translations and meaning norms. We also thank Heather Shrigley and Cassie Konkle for entering the similarity ratings, Ton Dijkstra, Béryl Schulpen, and Walter van Heuven for their helpful suggestions and comments on this project, and Peter Gianaros for comments on a previous version of this manuscript. Address correspondence to Natasha Tokowicz, Department of Psychology, 342C Baker Hall, Carnegie Mellon University, Pittsburgh, PA 15213 or by email at Tokowicz@cmu.edu.

Footnotes

1. These norms include a large number of measures that reflect the lexical and conceptual properties of these items. The complete list of the variables included in these norms will appear at http://www.academicpress.com/jml. The relation between these norms on form similarity, concreteness, and context availability are related to the number of translations, form similarity, and semantic similarity norms reported in the present study.

2. These data are available to view or download free of charge as a text file.

Table 1

Language History Questionnaire Data from the Participants of the Bilingual Norming Studies by Task

	Task	
Measure	Number of Translations Norms	Similarity Norms
Age (years)	21.6 (3.0)	21.2 (2.3)
Age Began L2 (years)	10.5 (2.1)	9.0 (2.5)
Time Studied L2 (years)	9.3 (3.6)	9.9 (3.8)
L2 Immersion Experience (months)	5.0 (12.3)	2.0 (4.0)
L2 Reading Ability	7.5 (1.3)	8.2 (.9)
L2 Writing Ability	6.5 (1.3)	7.1 (1.1)
L2 Conversation Ability	6.9 (1.2)	7.3 (1.3)
L2 Speech Comprehension Ability	8.0 (1.0)	8.2 (.9)

Note. Standard deviations are given in parentheses. Reading, writing, conversational, and speech comprehension ability were rated on a 10-point scale where 1 indicated the lowest level of ability and 10 indicated the highest level of ability.

Table 2

Number of Translations Data for the 562 Translation Pairs by Direction of Translation

Measure	Direction of Translation	
	Dutch to English	English to Dutch
Number of Translations	1.3 (.6)	1.4 (.8)
Number of Meanings Translated	1.1 (.3)	1.1 (.4)
Percent Expected Translations	74.2 (31.4)	75.3 (32.4)
Percent Synonym Translations	7.3 (19.3)	9.1 (21.7)
Percent Other Meaning Translations	3.6 (13.7)	4.8 (15.6)
Percent Verb Translations	0.1 (2.0)	0.5 (3.1)
Percent Colloquial Translations	0.3 (3.4)	0.0 (0)
Percent Unclassified Correct Translations	0.8 (8.1)	0.4 (2.8)
Percent Incorrect Translations	6.5 (13.2)	5.8 (13.1)
Percent Omitted Translations	7.1 (14.3)	4.2 (12.1)

Note. Standard deviations are given in parentheses.

Table 3

Intercorrelations among Factors

Factor	1	2	3	4	5	6	7	8
1. Semantic Similarity Rating	--	.02	-.12*	-.26**	.18**	.22**	.29**	.34**
2. Form Similarity Rating		--	-.10*	-.11*	.16**	.19**	.13**	.14**
3. Number of Translations Dutch to English			--	.14**	-.16**	-.22**	-.16**	-.16**
4. Number of Translations English to Dutch				--	-.23**	-.23**	-.26**	-.21**
5. Dutch Concreteness					--	.94**	.82**	.71**
6. English Concreteness						--	.80**	.80**
7. Dutch Context Availability							--	.77**
8. English Context Availability								--

Note. * $p < .05$, ** $p < .01$

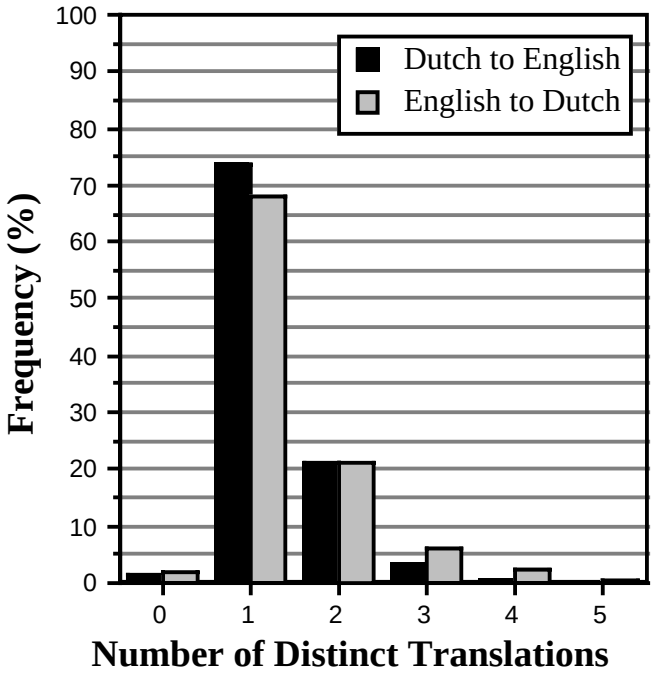
Figure Captions

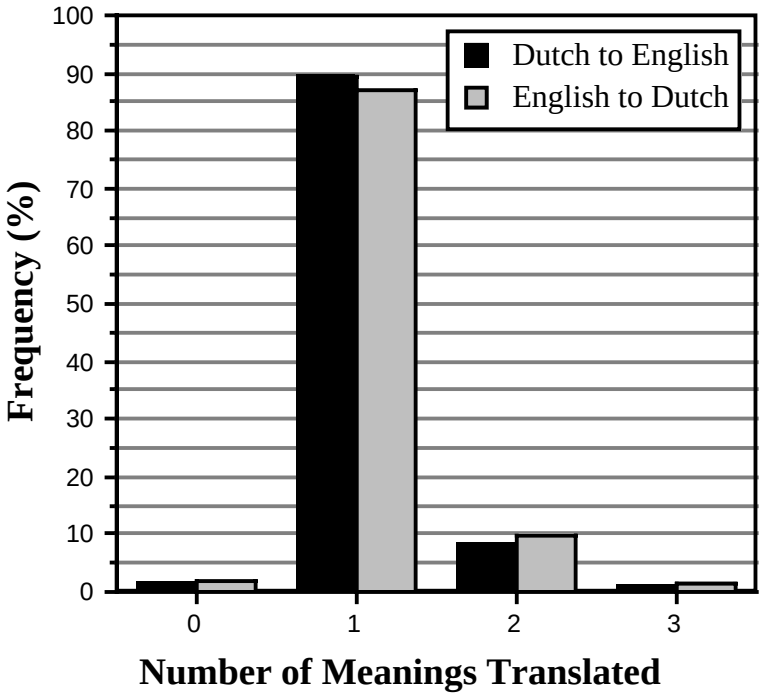
Figure 1. Histogram of number of translations for words translated from Dutch to English and English to Dutch.

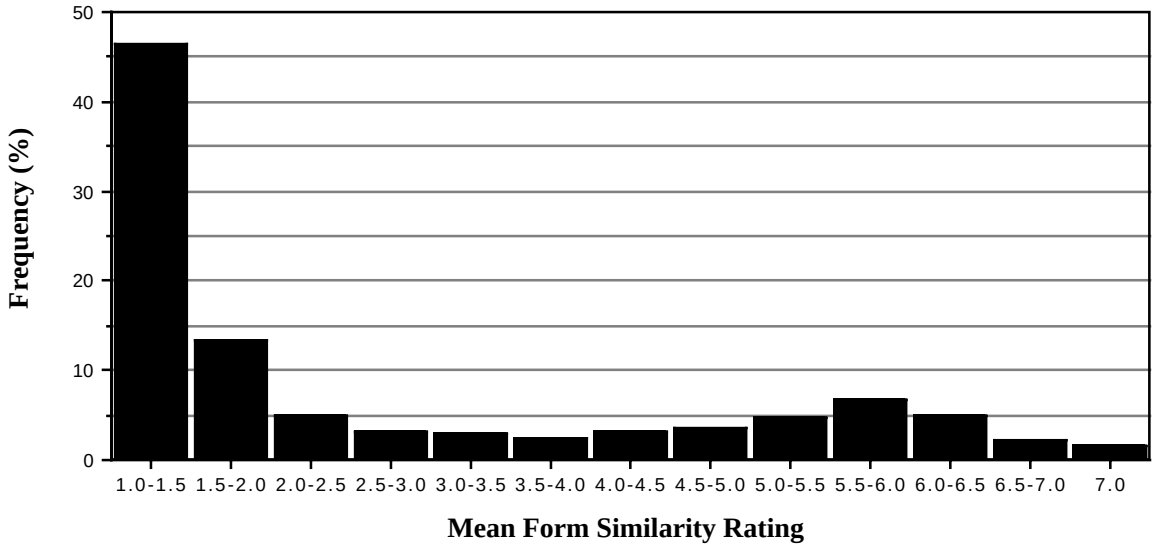
Figure 2. Histogram of number of meanings translated from Dutch to English and English to Dutch.

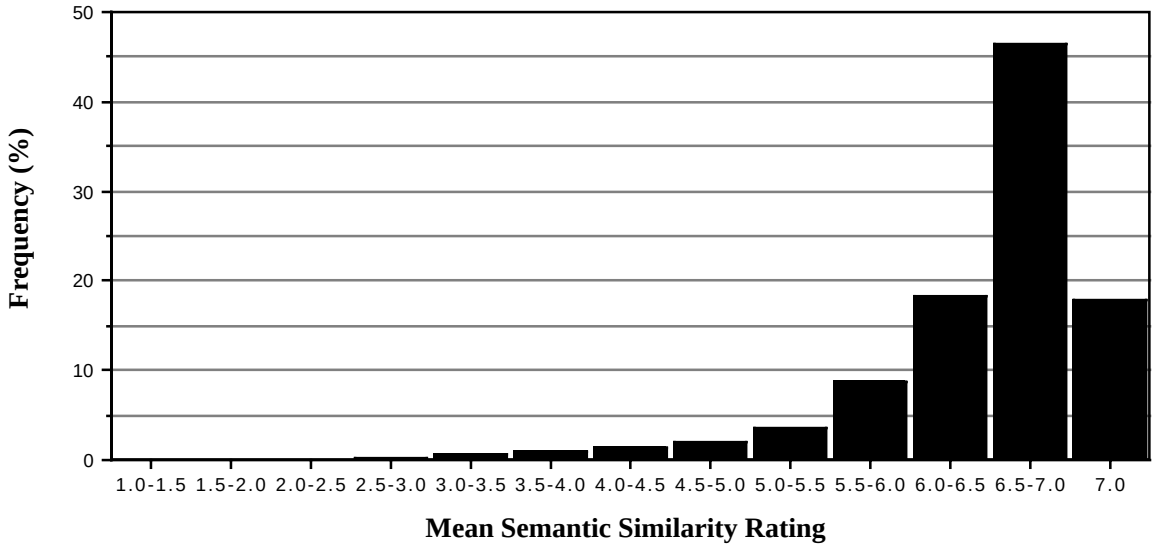
Figure 3. Histogram of mean form similarity ratings for all translation pairs.

Figure 4. Histogram of mean semantic similarity ratings for all translation pairs.


Dutch to English
English to Dutch
Frequency (%)
100
90
80
70
60
50
40
30
20
10
0
0
1
2
3
4
5
Number of Distinct Translations


100
90
80
70
60
50
40
30
20
10
0
Frequency (%)
Dutch to English
English to Dutch
0
1
2
3
Number of Meanings Translated


50
40
30
20
10
0
Frequency (%)
1.0-1.5
1.5-2.0
2.0-2.5
2.5-3.0
3.0-3.5
3.5-4.0
4.0-4.5
4.5-5.0
5.0-5.5
5.5-6.0
6.0-6.5
6.5-7.0
7.0
Mean Form Similarity Rating


Frequency (%)
50
40
30
20
10
0
1.0-1.5
1.5-2.0
2.0-2.5
2.5-3.0
3.0-3.5
3.5-4.0
4.0-4.5
4.5-5.0
5.0-5.5
5.5-6.0
6.0-6.5
6.5-7.0
7.0
Mean Semantic Similarity Rating