Behavior Research Methods, Instruments, & Computers
2002, 34 (3), 435-451

Number-of-translation norms for Dutch–English translation pairs: A new tool for examining language production

NATASHA TOKOWICZ
Carnegie Mellon University, Pittsburgh, Pennsylvania

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

ANNETTE M. B. DE GROOT
University of Amsterdam, Amsterdam, The Netherlands

and

JANET G. VAN HELL
University of Nijmegen, Nijmegen, The Netherlands

We collected number-of-translation 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 1,003 possible translation pairs. Approximately 40% of the translations were rated as being similar across languages with respect to spelling/sound (i.e., they were cognates). Approximately 45% of the translations were rated as being highly semantically similar across languages. At least 25% of the words in each direction of translation had more than one translation. The form similarity ratings were found to be highly reliable even when obtained with different bilinguals and modified rating procedures. Number of translations and meaning factors significantly predicted the 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. These norms are available at http://www.talkbank.org/norms/tokowicz/.

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 for 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, and Steehouwer (1993). In this task, participants are given a definition, such as *an animal that barks*, and respond by naming the object to which the definition refers (in this example, *dog*). 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 Eckhardt, and

This research was supported in part by a teaching release grant to N.T. from the Research and Graduate Studies Office at 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 J.F.K. The norming data were collected as part of the doctoral thesis awarded to the first author from 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 and Ton Dijkstra, Erica Michael, Béryl Schulpen, and Walter van Heuven for their useful suggestions and comments on this project. We thank Marc Brysbaert, Peter Gianaros, Nan Jiang, and Jonathan Vaughan for helpful comments on a previous version of the manuscript. Address correspondence to J. F. Kroll, Professor of Psychology and Applied Linguistics, 641 Moore Building, Department of Psychology, Pennsylvania State University, University Park, PA 16802 (e-mail: tokowicz@alumni.umass.edu).

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) later showed that the translation task also could 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, in a series of studies, the processes underlying translation production have been examined without direct comparison with 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, 2002; 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, 2002). 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 (e.g., word frequency, concreteness, ambiguity). 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 previous 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) for the study of 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 does not 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., owing 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 the 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, 2002; 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 1,003 possible translation pairs. These translation pairs include the 562 original translation pairs from the previous studies, as well as the additional translation pairs that were generated by the norming participants (i.e., we paired each word with each of its possible translations). Note that although some of the translations that were expected in the original studies were never given by the norming participants, we obtained similarity ratings on these translation pairs. Thus, these norms allowed us 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 are 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 that are 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. (2002), 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 the semantic similarity of translation equivalents, form overlap has been shown to affect bilingual language performance on a variety of tasks, including translation (e.g., Kroll & Stewart, 1994), word naming (e.g., Schwartz, Kroll, & Diaz, 2001), picture naming (e.g., Costa, Caramazza, & Sebastián-Gallés, 2000), and lexical decision (e.g., de Groot & Nas, 1991). 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; see, 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, naive 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, 2001, for a comparison of several cognate identification techniques, including cognate norms collected from naive 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 may be affected by their 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., 2002; de Groot et al., 1994; Dijkstra et al., 1999; van Hell, 1998; van Hell & de Groot, 1998a, 1998b). Subsequently, all possible translation pairs that had been used in the previous studies and that were generated from the number-of-translations norming study (i.e., all words paired with all other possible translations) were rated on form and semantic similarity. These norms are available at http://www.talkbank.org/norms/tokowicz/.[2] See Table 1 for a list of the variables for which data are available on the Web page.

Table 1
Variables for Which Data Are Available on the Norm Web Page

Variable
1. Length of Dutch and English words (number of letters)
2. Number of translations from Dutch to English and English to Dutch
3. Number of meanings translated from Dutch to English and English to Dutch
4. Form similarity rating of the translation pair (average across directions and separated by direction)*
5. Semantic similarity rating of the translation pair (average across directions and separated by direction)*
6. First through fourth alternate translations from Dutch to English
7. Length of first through fourth alternate translations from Dutch to English
8. Form similarity rating of Dutch stimulus and first through fourth alternate translations
9. Semantic similarity rating of Dutch stimulus and first through fourth alternate translations
10. First through fifth alternate translations from English to Dutch
11. Length of first through fifth alternate translations from English to Dutch
12. Form similarity rating of English stimulus and first through fifth alternate translations
13. Semantic similarity rating of English stimulus and first through fifth alternate translations
14. Source of stimulus (de Groot et al., 2002, or Dijkstra, Grainger, and Van Heuven, 1999)
15. Form similarity rating of original translation pair from Dutch to English and English to Dutch
16. Semantic similarity rating of original translation pair from Dutch to English and English to Dutch
17. Percentage of expected translations from Dutch to English and English to Dutch
18. Percentage of synonym translations from Dutch to English and English to Dutch
19. Percentage of other meaning translations from Dutch to English and English to Dutch
20. Percentage of verb meaning translations from Dutch to English and English to Dutch
21. Percentage of colloquial use translations from Dutch to English and English to Dutch
22. Percentage of unclassified correct translations from Dutch to English and English to Dutch
23. Percentage of erroneous translations from Dutch to English and English to Dutch
24. Percentage of omitted translations from Dutch to English and English to Dutch

*The form and semantic similarity ratings, separated by direction of translation, are available in a separate file (see the Web page for instructions).

Number-of-Translation Norms

Number-of-translation 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. For example, when *trunk* was the stimulus word and *achterback* was given as the translation, we inferred that the car meaning of trunk had been translated. In contrast, when *slurf* was given as the translation, we inferred that the snout meaning of trunk had been translated. Note that although we refer to this derived measure as *number of meanings*, 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; there may be multiple meanings of a word that do not require more than one translation across languages.

There are many possible ways in which the number of possible translations of a word can be calculated. We chose to use the *first translation* method adopted by Schönpflug (1997) and others in the polysemy literature (e.g., Rubenstein, Garfield, & Millikan, 1970), because this method matches the characteristics of the translation production task on which we were trying to predict performance (see Tokowicz, 2000) and also seems to have the fewest limitations. One alternative to the procedure we selected is to ask each participant to write all of the translations they know for each word. However, it is unclear how best to extrapolate the number of translations from such a measure, since there are many options (the largest number of translations any individual gave, the total number of translations that were given by all the participants, the total number of translations given overall, etc.). Although the latter procedure may seem to be a more direct measure of the number of translations of which each bilingual is aware, we believe that the measure we used also approximates the number of translations known by the average bilinguals in our population. This is because translations that are powerful enough to affect performance should be listed at least once in a normative study of this type.

Participants. The participants in the number-of-translati on norming task were 24 Dutch–English bilinguals who were students at the University of Nijmegen. The participants completed language history questionnaires in which they reported on their L2 learning experiences. The participants rated their L1 and L2 reading, writing, conversational, and speech comprehension abilities on a scale that ranged from 1 to 10 and indicated the age at which L2 learning began and the types of exposure they had when learning the L2 (see Table 2 for the language history questionnaire data).

Procedure. The words were divided into several list versions and were 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 with 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. Note that because our stimuli had been used in previous experiments in which one translation had been considered the best translation, for the purposes of computing number of translations, we considered the best translation in the previous experiment the *expected translation* in our normative study.

The number of correct possible translations for each word was calculated. These correct translations came from several possible types of responses: expected translations, synonyms of the expected translation, other meanings (when a different meaning of the stimulus word was translated), verb meanings, and colloquial uses. We calculated 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.

Form Similarity and Semantic Similarity Norms

Ratings of form similarity[3] and semantic similarity were obtained for all correct translation pairs that had been given in the number-of-translation 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 2 for the language history questionnaire data). The participants in the similarity rating study had not participated in the number-of-translat ion norming study.

Procedure. The 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[4]; 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 Appendix A for the rating instructions). The instructions indicated that the participants were to complete the semantic similarity rating for a pair first 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 having difficulty rating the word pairs with respect to their spelling and sound similarity without taking their semantic similarity into account. By performing the semantic similarity rating first, the participants could emphasize form similarity in the second rating.

Table 2
Language History Questionnaire Data From the Participants of the Bilingual Norming Studies by Task

	Task			
	Number of Translation Norms		Similarity Norms	
Measure	*M*	*SD*	*M*	*SD*
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	0.9
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	0.9

Note—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 3
Number-of-Translations Data for the 562 Translation Pairs by Direction of Translation

	Direction of Translation			
	Dutch to English		English to Dutch	
Measure	*M*	*SD*	*M*	*SD*
Number of translations	1.3	0.6	1.4	0.8
Number of meanings translated	1.1	0.3	1.1	0.4
Percentage of expected translation	74.2	31.4	75.3	32.4
Percentage of synonym translations	7.3	19.3	9.1	21.7
Percentage of other meaning translations	3.6	13.7	4.8	15.6
Percentage of verb translations	0.1	2.0	0.5	3.1
Percentage of colloquial translations	0.3	3.4	0.0	0.0
Percentage of unclassified correct translations	0.8	8.1	0.4	2.8
Percentage of incorrect translations	6.5	13.2	5.8	13.1
Percentage of omitted translations	7.1	14.3	4.2	12.1

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 3). Most of the words were assigned only one translation (see Figure 1). In both directions, the range was from zero (i.e., an acceptable translation was never given) to five. At least 20% of the 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 from Dutch to English [$M_{\text{Dutch to English}} = 1.29$, $SD = 0.58$; $M_{\text{English to Dutch}} = 1.41$, $SD = 0.78$; $t(561) = 3.2$, $p < .01$], which may reflect the participants' larger vocabulary in Dutch. Alternatively, it is possible that Dutch is a more polysemous language than English and, therefore, more translations are possible in the English to Dutch direction of translation (see, e.g., the difference between English and Italian suggested by Hunt & Agnoli, 1991).

The number of meanings that were translated from Dutch to English and from 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 10% 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 examined.

Form Similarity and Semantic Similarity Ratings

The rating results reflect the data from the set of all 1,003 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). See Appendix B for the form and semantic similarity ratings for all word pairs.

Although the English to Dutch similarity ratings were slightly higher than the Dutch to English ratings [$M_{\text{English to Dutch}} = 2.7$, $SD = 2.0$; $M_{\text{Dutch to English}} = 2.6$, $SD = 1.9$; $t(1,002) = 7.8$, $p < .01$], the correlation between them was very high ($r = .95$, $p < .01$). 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 ($M = 6.4$, $SD = 0.7$; see Figure 4 for the distribution of the mean semantic similarity ratings). Furthermore, no word pair received a mean rating lower than 2.5. As for the form sim-

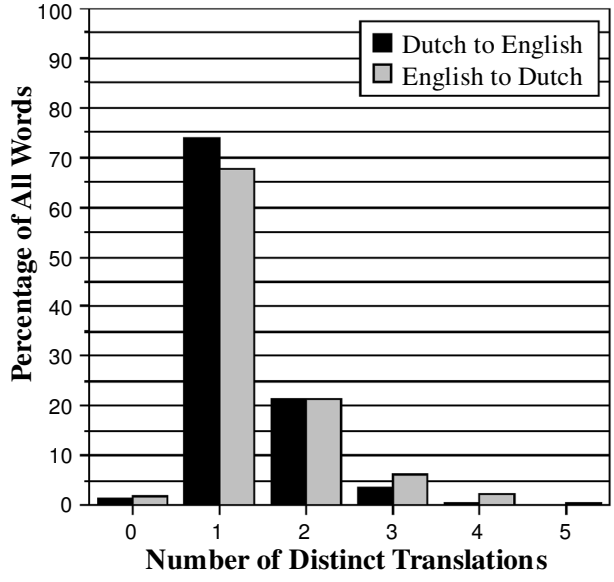

Figure 1. Percentage of number of translations for words translated from Dutch to English and from English to Dutch.

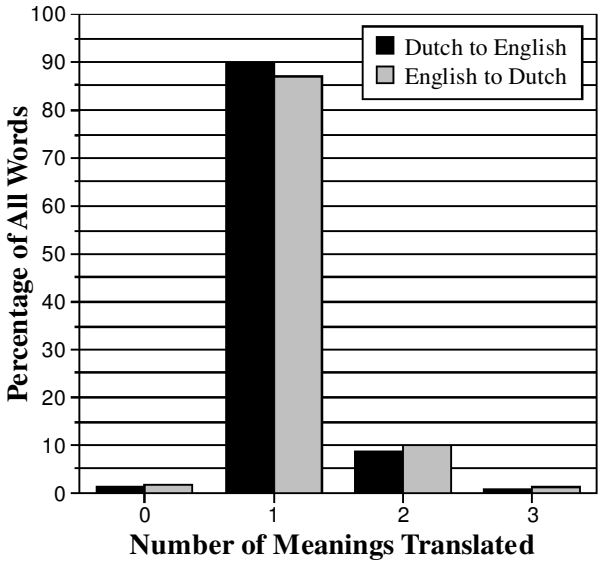

Figure 2. Percentage of number of meanings translated from Dutch to English and from English to Dutch.

ilarity 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 [$M_{\text{English to Dutch}} = 6.44$, $SD = 0.76$, $M_{\text{Dutch to English}} = 6.36$, $SD = 0.83$; $t(1{,}002) = 3.7$, $p < .01$]. Nevertheless, 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 cross-language form overlap norms (combined spelling and sound) from a group of students at the University of Amsterdam on a set of Dutch–English translation pairs, a subset of which had been rated previously in a study by de Groot and Nas (1991). De Groot reported that the words for which norms had been obtained in both studies were highly correlated ($r = .98$). We obtained form similarity norms for a large set of words, a subset of which were the 440 translation pairs 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 10 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., 2002). The factors included in this correlational analysis are form similarity, number of translations, concreteness (from de Groot et al., 2002), context availability (from de Groot et al., 2002), and semantic similarity. These correlations were performed using the mean form and semantic similarity ratings across languages (see Table 4).

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 semantically similar to any one of their translations than are words with only a single translation. Indeed, semantic similarity is significantly negatively correlated with the number of distinct translations a word has, indicating that the higher the number of translations, the

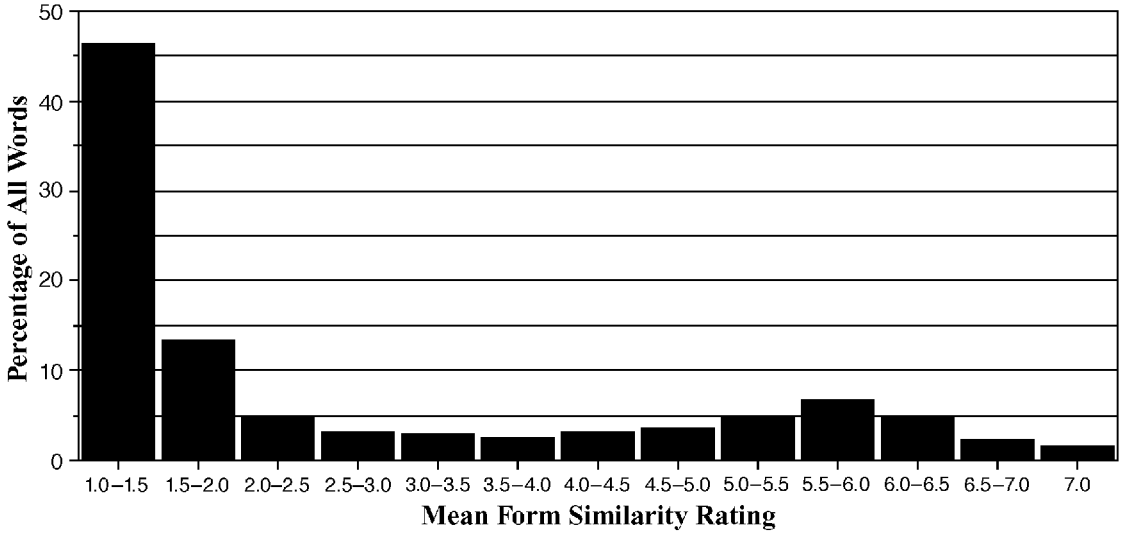

Figure 3. Percentage of mean form similarity ratings for all translation pairs.

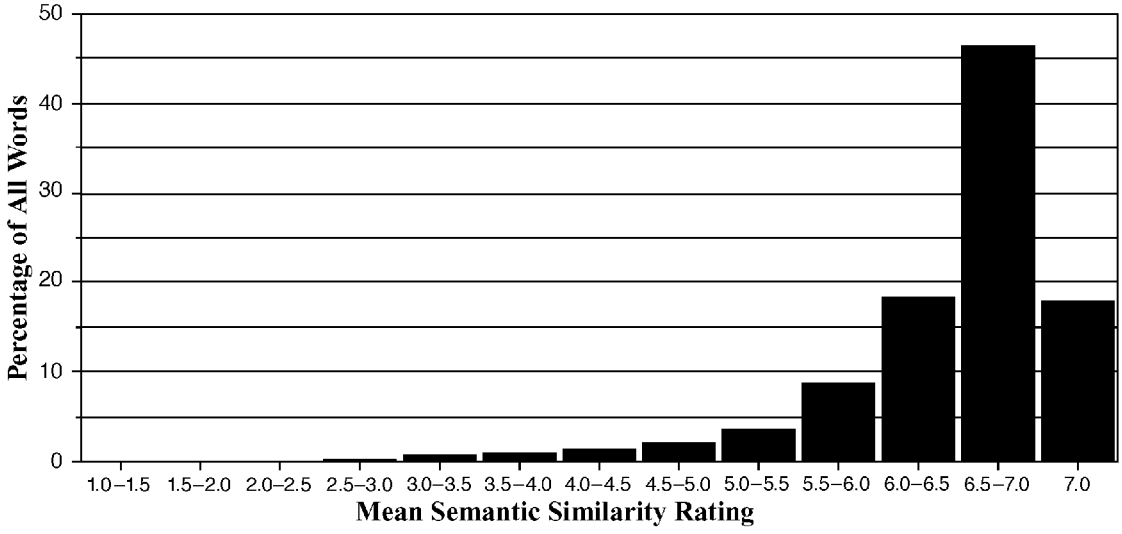

Figure 4. Percentage of mean semantic similarity ratings for all translation pairs.

lower the similarity of the translation pair with respect to meaning. In fact, studies in which the effects of number of translations on translation performance have been examined 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, 2002). 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. Alternatively, multiple-translation words may be less semantically similar than single-translation words, which results in slower and less accurate translation.

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 are words low in concreteness. The data are consistent with this suggestion; semantic similarity is correlated with word concreteness, indicating that the higher the concreteness of the word, the higher the meaning similarity of that word to its translation. Context availability reflects the ease with which context can be accessed for a word and is typically highly correlated with concreteness, indicating that words higher in concreteness are also higher in their context availability (rs = .73 and .65, Schwanenflugel, Harnishfeger, & Stowe, 1988, Experiments 2 and 3, respectively; r = .88, Schwanenflugel & Shoben, 1983). In the de Groot et al. (2002) norms, concreteness and context availability are also significantly correlated. Furthermore, context availability is correlated with semantic similarity, indicating that the higher the context availability, the more semantically similar the word is to its translation.

In our sample of words, semantic similarity did not correlate significantly with form similarity across languages.[5] This finding is not consistent with de Groot's (1992) hypothesis that words that 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 con-

Table 4
Intercorrelations Among Factors

Factor	1	2	3	4	5	6	7	8
1. Semantic similarity rating	–	.02	−.12*	−.27**	.18**	.22**	.29**	.34**
2. Form similarity rating		–	−.10*	−.12*	.16**	.19**	.13**	.14**
3. Number of translations from Dutch to English			–	.14**	−.16**	−.22**	−.16**	−.16**
4. Number of translations from English to Dutch				–	−.24**	−.24**	−.27**	−.21**
5. Dutch concreteness					–	.94**	.82**	.71**
6. English concreteness						–	.80**	.80**
7. Dutch context availability							–	−.77**
8. English context availability								–

*p < .05. **p < .01.

creteness/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 from 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, indicating that higher concreteness/context availability was associated with higher semantic similarity [$R^2 = .14$; $F(4,433) = 17.7$, $p < .01$]. Because context availability and concreteness are significantly correlated in our sample, it is not possible to examine the effects of these factors independently because they would account for largely the same proportion of variance. Therefore, we cannot determine whether concreteness of context availability is a better predictor of semantic similarity.

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 [$\Delta R^2 = .04$; $F\Delta(2,430) = 11.5$, $p < .01$]. Again, due to the significant correlation between number of translations from English to Dutch and Dutch to English, it is not possible to determine which is a better predictor of semantic similarity. Thus, the results from the hierarchical regression analysis were consistent with the conclusions drawn on the basis of the previous correlations. Specifically, concreteness/context availability and number of translations accounted for unique proportions of variance in semantic similarity ratings, whereas form similarity did not.

CONCLUSIONS

Like the information available regarding name agreement for pictured objects, the present set of norms makes 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-translation norms on a set of 562 Dutch–English translation pairs. We then obtained form and semantic similarity ratings on the set of 1,003 possible Dutch–English translation pairs that came from the original studies and that were generated during the number-of-translation norming task. Overall, the majority of the words had only a single translation across languages both from Dutch to English and from 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 were highly correlated with those obtained in past norming studies. Intercorrelations among the factors showed that semantic similarity was correlated with concreteness and context availability, so 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 was not correlated with form similarity, at least in the sample of words used in this study. These correlational results were confirmed with 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 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behavior Research Methods, Instruments, & Computers*, **31**, 531-552.

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.

DE GROOT, A. M. B. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 1001-1018.

DE GROOT, A. M. B., BORGWALDT, S., BOS, M., & VAN DEN EIJNDEN, E. (2002). Lexical decision and word naming in bilinguals: Language effects and task effects. *Journal of Memory & Language*, **47**, 91-124.

DE GROOT, A. M. B., DANNENBURG, L., & VAN HELL, J. G. (1994). Forward and backward word translation by bilinguals. *Journal of Memory & 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 & 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 & Language*, **41**, 496-518.

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.

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

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

HUNT, E., & AGNOLI, F. (1991). The Whorfian hypothesis: A cognitive psychology perspective. *Psychological Review*, **98**, 377-389.

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, & Cognition*, **24**, 1256-1274.

KROLL, J. F., MICHAEL, E. B., TOKOWICZ, N., & DUFOUR, R. (2002). The development of lexical fluency in a second language. *Second Language Research*, **18**, 137-171.

KROLL, J. F., & STEWART, E. (1994). Concept 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. (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.

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, & 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 & Brain Sciences*, **22**, 1-75.

Pavlenko, A. (1999). New approaches to concepts in bilingual memory. *Bilingualism: Language & 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, & 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 & Verbal Behavior*, **23**, 23-38.

Prisma Woordenboek: Engels–Nederlands (1990). Utrecht: Uitgeverij Het Spectrum.

Prisma Woordenboek: Nederlands–Engels (1990). Utrecht: Uitgeverij Het Spectrum.

Rubenstein, H., Garfield, L., & Millikan, J. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning & Verbal Behavior*, **9**, 487-494.

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, U.K.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory & 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, & Cognition*, **9**, 82-102.

Schwartz, A., Kroll, J. F., & Diaz, M. (2001, November). *Mapping orthography to phonology: More evidence for nonselective cross-language activation*. Poster presented at the 42nd Annual Meeting of the Psychonomic Society, Orlando, FL.

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, & Cognition*, **21**, 686-698.

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

Tokowicz, N., & Kroll, J. F. (2002). *Speaking words in two languages: The role of multiple translations in bilingual production*. Manuscript in preparation.

van Hell, J. G. (1998). *Cross-language processing and bilingual memory organization*. Unpublished doctoral 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 & 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*, **51A**, 41-63.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition*, **15**, 181-198.

NOTES

1. These norms include a large number of measures that reflect the lexical and conceptual properties of the words. The complete list of the variables included in these norms will appear at http://www.academicpress.com/jml. In the present paper, we relate our measures of form similarity, semantic similarity, and number of translations to de Groot et al.'s (2002) form similarity, concreteness, and context availability norms.

2. These data are available to view or download free of charge.

3. We chose to use a subjective measure of the form overlap of translation equivalents, rather than objective measures (e.g., Van Orden, 1987), because we were most interested in the effects of the perceived similarity by proficient bilinguals on the translation performance of a similar group of bilinguals. Our results demonstrate that this measure indeed predicts performance on the task, in such a way that translations high in form similarity are translated more quickly and accurately than translations that are low in form similarity (Tokowicz, 2000).

4. Some researchers have examined the independent consequences of shared orthography and phonology across languages. In a progressive demasking task and a visual lexical decision task in which bilingual participants were to respond to words in their L2, Dijkstra et al. (1999) found that cognates were responded to more quickly than noncognates only when they shared orthography (or orthography and phonology) across languages, but not when they shared phonology only. Although this result may suggest that our use of a combined spelling and sound measure limits our ability to predict performance in bilingual tasks, this will clearly depend on the nature of the task of interest. In the present study, the goal was to predict bilingual translation production performance. In the translation production task, unlike the receptive language tasks, shared phonology may facilitate performance because the participants respond by speaking words aloud. Note also that unlike the tasks used by Dijkstra et al., our task necessitated the simultaneous activation of L1 and L2.

To our knowledge, the independent consequences of orthographic and phonological overlap on bilingual language production have been examined in only one study. Schwartz et al. (2001) found that bilinguals were faster to translate or read aloud cognates that shared both orthography and phonology than to translate or read aloud cognates that shared only orthography or phonology. Thus, the measure we used may be sufficient to predict performance on the task of interest, because we used a measure that weights spelling and sound equally; translations that overlap only on one dimension will be rated as lower in their overlap than will translations that overlap on both dimensions. Note that the measure we obtained had been used in the past to predict translation performance (de Groot & Nas, 1991) and was a significant predictor of translation time and accuracy in a translation production task (Tokowicz, 2000).

5. Note that certain types of words, such as scientific terms, may be underrepresented in our sample. Therefore, without further study, it is not possible for us to generalize our null correlation between form and semantic similarity to these types of words. We thank Marc Brysbaert for calling our attention to this bias.

APPENDIX A
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." *Your rating should reflect a combination of both the spelling and sound similarity.*

		Meaning							Spelling/Sound						
Examples		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**
citroen	arrival	**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.

APPENDIX B
Form and Semantic Similarity Ratings for All Word Pairs

Dutch Word	English Word	Form Similarity	Semantic Similarity	Dutch Word	English Word	Form Similarity	Semantic Similarity
aandacht	attention	2.62	6.88	bedrieging	betrayal	2.38	6.00
aankomst	arrival	2.38	7.00	bedrog	betrayal	3.38	5.88
aanmelding	entry	1.38	6.00	beeld	insight	1.12	5.38
aanzoek	proposal	1.25	6.50	been	leg	2.00	6.88
aap	ape	6.12	5.88	begrafenis	funeral	1.12	6.88
aap	monkey	1.12	7.00	bekentenis	confession	1.12	7.00
aard	nature	1.12	6.38	belangstelling	interest	1.25	6.50
aardappel	potato	1.00	6.88	belofte	promise	1.25	6.75
aardbei	strawberry	1.00	7.00	benodigdheid	necessity	1.75	6.00
aarde	earth	5.12	7.00	benzine	gas	1.00	6.00
aarde	soil	1.25	5.88	berg	mountain	1.38	7.00
aardig	nice	1.12	6.62	bericht	message	1.00	6.88
achterbak	trunk	1.25	6.75	beroemd	fame	1.12	5.62
actie	action	5.75	6.88	bescherming	protection	1.25	6.88
adelaar	eagle	2.00	6.29	beschrijving	description	1.75	6.88
afkeer	dislike	1.25	6.12	beslissing	decision	1.38	6.88
afstand	distance	1.25	7.00	besluit	decision	1.43	6.86
afval	rubbish	1.25	5.88	bestelling	order	1.38	6.50
afval	trash	1.43	6.29	betekenis	meaning	1.12	7.00
afval	waste	1.12	6.00	bevel	command	1.12	6.62
appel	apple	6.62	7.00	bevel	order	1.75	6.50
arend	eagle	2.75	7.00	bewijs	evidence	1.38	6.38
attentie	attention	6.00	6.75	bewijs	proof	1.29	6.86
auteur	author	6.12	7.00	bewijs	prove	1.25	6.38
auteur	writer	1.38	5.88	bezit	possession	1.25	6.62
auto	car	1.50	6.75	bezit	property	1.00	6.25
bad	bath	5.50	6.88	biecht	confession	1.25	6.75
bakker	baker	5.62	7.00	bier	ale	1.25	5.62
bakker	bakery	5.12	5.88	bier	beer	6.12	6.88
bal	ball	6.38	7.00	bij	at	1.25	5.88
bal	dance	1.29	5.57	bij	bee	4.38	7.00
bandiet	crook	1.12	6.62	bij	with	1.12	4.88
bank	bank	6.88	5.88	bijbel	bible	4.88	7.00
bank	bench	4.00	5.62	bijeenkomst	meeting	1.38	6.75
bank	couch	1.25	6.75	blaam	blame	5.75	6.50
bankje	bench	4.25	6.29	blad	sheet	1.14	6.25
basis	base	4.62	6.38	bleek	fair	1.38	3.00
basis	basic	4.88	5.75	bliksem	lightning	1.88	6.75
basis	basis	7.00	6.75	blind	blind	6.75	6.88
bedoeling	meaning	1.38	6.25	bloem	flower	1.62	6.88
bedreiging	threat	1.00	6.62	blok	block	6.12	6.62
bedrieger	crook	1.12	5.75	blok	square	1.12	4.12

APPENDIX B (Continued)

Dutch Word	English Word	Form Similarity	Semantic Similarity
blokkade	block	4.12	4.75
bod	bid	5.62	6.62
bod	offer	1.00	6.00
boek	book	6.25	7.00
boer	farmer	1.12	6.88
boerderij	farm	1.12	7.00
boezem	bosom	5.43	6.71
boezem	breast	2.50	5.75
bont	fur	1.25	6.62
boodschap	errand	1.38	6.00
boodschap	message	1.38	6.88
boom	tree	1.29	6.86
boomstronk	trunk	2.38	5.25
boosheid	anger	1.25	6.38
boot	boat	6.12	6.88
boot	ship	1.12	6.50
borst	breast	5.00	6.38
borst	chest	2.38	6.43
borstkas	chest	2.12	6.50
bos	forest	1.50	6.75
bos	wood	1.25	5.62
bot	blunt	2.12	5.50
bot	bone	3.12	6.88
bot	rude	1.00	6.38
broek	pants	1.12	6.75
broek	trousers	1.12	6.50
broer	brother	3.71	7.00
brood	bread	5.62	6.62
bruid	bride	5.38	7.00
bureau	desk	1.12	6.38
cadeau	gift	1.25	6.88
cape	cape	7.00	3.50
cape	cloak	2.75	5.38
cirkel	circle	6.38	6.88
cirkel	cycle	3.25	2.88
citroen	lemon	1.43	6.43
creatie	creation	5.88	6.38
crimineel	crook	2.00	6.62
cru	crude	5.75	6.25
cultuur	culture	5.75	7.00
cyclus	circle	3.38	3.25
cyclus	cycle	5.38	6.12
daad	deed	5.50	6.12
dag	day	6.12	7.00
dageraad	dawn	2.50	6.75
dak	roof	1.12	6.88
daling	decline	2.62	6.12
daling	descent	2.50	6.38
daling	fall	1.12	4.38
dame	dame	7.00	5.71
dame	lady	1.25	6.62
dans	dance	6.25	6.62
dansen	dance	4.75	7.00
dapper	bold	1.12	6.12
dapper	brave	1.38	6.62
darm	bowel	1.14	4.29
datum	date	5.00	7.00
deel	part	1.62	6.25
deken	blanket	1.43	6.00
den	pine	1.12	6.88
deugd	virtue	1.50	6.88
dief	crook	1.62	5.88
dief	thief	5.25	6.88
dienstmeisje	maid	1.88	6.75
dierenarts	vet	1.12	7.00
dierenarts	veterinarian	1.25	6.88
dij	thigh	5.29	7.00
ding	thing	6.17	6.86
directie	direction	5.86	3.29
dochter	daughter	4.12	6.88
dokter	doctor	6.25	6.88
domein	domain	6.00	7.00
domein	property	1.38	4.12
doos	box	1.75	7.00
dorp	village	1.12	6.88
dorpje	village	1.38	7.00
draad	thread	3.62	6.00
draad	wire	1.25	6.50
draadje	thread	2.38	6.25
dreigement	threat	1.62	6.62
dreiging	threat	2.50	6.50
driehoek	triangle	1.62	7.00
druif	grape	1.62	6.12
drukte	crowd	1.62	4.75
duif	dove	4.38	6.88
duif	pigeon	1.12	6.75
duim	thumb	3.00	7.00
dun	slim	1.57	5.86
dun	thin	2.88	6.88
dwang	compulsion	1.38	6.38
dwang	pressure	1.12	6.12
ede	oath	2.75	6.75
edel	noble	1.62	6.75
educatie	education	6.12	6.50
eed	oath	3.50	6.62
eend	duck	1.50	7.00
eenheid	measure	1.50	5.25
eenheid	unity	2.75	6.38
eenvoud	simplicity	1.12	6.88
eenvoud	single	1.25	3.25
eenvoudig	simplicity	1.00	6.25
eenvoudigheid	simplicity	1.25	6.50
eerbied	honor	1.25	6.38
eerbied	respect	1.50	6.14
eerlijk	fair	1.25	6.38
eerlijk	honest	1.38	7.00
eerlijkheid	honesty	1.38	6.75
eeuw	century	1.12	6.88
eigenaar	owner	2.25	6.88
eigenschap	virtue	1.14	5.00
eind	end	5.50	6.88
einde	end	5.88	7.00
eis	demand	1.12	6.38
elementair	crude	1.12	2.50
elleboog	elbow	5.62	7.00
enorm	gigantic	1.62	6.12
enorm	huge	1.25	6.88
enorm	tremendous	1.75	6.62
entree	entry	5.88	6.12
erfenis	heritage	1.75	6.62
erfenis	inheritance	1.75	6.75
ervaring	experience	2.25	6.75
erwt	pea	1.14	6.29
evenement	event	3.25	6.00
fabriek	mill	1.00	4.25
fakkel	torch	1.25	7.00
fiets	bicycle	1.00	6.75
fiets	bike	1.50	6.38
fiets	cycle	1.38	5.88
figuur	figure	5.71	5.00
film	film	6.88	6.75
film	movie	1.25	7.00
fles	bottle	1.62	7.00
formulier	form	4.62	6.62
fruit	fruit	6.75	6.62
gala	ball	1.38	6.38

APPENDIX B (Continued)

Dutch Word	English Word	Form Similarity	Semantic Similarity	Dutch Word	English Word	Form Similarity	Semantic Similarity
gangetje	alley	1.25	5.50	hals	throat	1.38	3.88
gas	gas	7.00	6.75	hand	hand	7.00	6.88
gast	chap	1.88	5.50	handdoek	towel	1.25	7.00
gat	gap	5.00	5.88	handel	trade	1.38	6.38
gazon	lawn	1.62	6.62	handschoen	glove	1.00	7.00
gebeurtenis	event	1.25	7.00	handtekening	autograph	1.12	5.88
gebeurtenis	happening	1.12	6.86	handtekening	signature	1.12	7.00
gebrek	deficiency	1.25	6.75	hard	tough	1.38	5.75
gebrek	flaw	1.00	5.88	hardheid	cruelty	1.00	5.25
gebrek	lack	2.12	6.88	hart	heart	5.00	6.88
gebruiksvoorwerp	utensil	1.25	6.50	hartstocht	passion	1.38	6.88
gedachte	thought	2.88	6.88	haven	haven	6.88	4.50
geduld	patience	1.29	6.86	haver	oat	1.00	6.75
geheugen	memory	1.00	6.88	heer	gentleman	1.00	6.62
geld	money	1.00	7.00	heer	sir	1.38	6.38
gelegenheid	occasion	1.25	6.50	heg	hedge	4.50	6.00
gelegenheid	opportunity	1.38	6.25	heiden	heathen	3.71	6.00
geloof	believe	1.25	6.50	heiden	pagan	1.43	5.57
geloof	faith	1.50	6.38	heilig	holy	4.25	7.00
geloof	religion	1.38	7.00	heilig	sacred	1.38	6.62
geluid	noise	1.12	5.50	hel	hell	6.62	7.00
geluk	happiness	1.12	6.88	held	hero	4.88	7.00
geluk	luck	3.75	6.75	hemel	sky	1.75	5.62
gemak	ease	1.38	6.62	herfst	autumn	1.12	6.88
gemeen	crude	1.25	6.38	herfst	fall	1.25	6.62
gemeen	cruel	1.25	6.75	herinnering	memory	1.50	6.12
genade	mercy	1.38	6.62	herstel	recovery	1.75	6.50
genade	pardon	1.12	4.50	herstellen	recovery	2.12	6.25
gerucht	rumour	2.62	6.88	hertog	duke	1.12	6.62
geschenk	gift	1.62	6.62	hoed	cap	1.12	5.25
geur	smell	1.50	6.38	hoed	hat	5.50	6.75
gevaar	danger	1.25	7.00	hoek	angle	1.12	6.88
geval	case	1.50	6.38	hoek	corner	1.12	6.88
gevangenis	jail	1.12	7.00	hol	cave	1.38	6.00
gevangenis	prison	1.00	7.00	hol	hollow	5.25	6.12
gevoel	feeling	2.88	6.50	hond	dog	1.75	7.00
geweer	gun	2.50	6.50	honing	honey	5.25	6.88
geweer	rifle	1.00	6.62	hoofd	head	4.25	6.88
geweten	conscience	1.00	6.75	hoofd	master	1.12	4.50
gezicht	face	1.12	6.88	hoogte	height	4.62	5.75
gezondheid	health	1.50	7.00	horloge	watch	1.12	6.88
gift	gift	6.88	5.25	hout	wood	1.75	6.62
gigantisch	huge	1.12	6.88	huid	skin	1.12	6.75
gitaar	guitar	6.00	7.00	huidig	current	1.25	6.12
golf	golf	6.88	7.00	huis	house	5.38	7.00
golf	wave	1.75	6.50	huishoudhulp	maid	1.00	5.75
goud	gold	5.62	7.00	huurder	renter	1.75	6.75
graaf	duke	1.00	6.75	huurder	tenant	1.14	6.57
graan	corn	1.38	6.12	ijzer	iron	3.12	6.62
grap	joke	1.12	6.75	impressie	impression	5.88	6.75
grapje	joke	1.25	6.88	indruk	impression	2.25	6.62
grappig	funny	1.38	6.75	influentie	influence	6.12	4.88
grasveld	lawn	1.12	6.50	informatie	information	5.62	7.00
griep	flu	1.00	7.00	ingang	entrance	3.25	6.88
groenteboer	greengrocer	3.25	6.50	ingang	entry	1.75	6.75
groenteman	greengrocer	3.88	6.25	inhoud	content	1.38	5.88
grond	floor	1.43	5.43	inhoud	contents	1.43	6.71
grond	ground	5.88	6.75	inhoud	volume	1.25	5.62
grond	soil	1.12	6.50	inkt	ink	6.12	7.00
grondbeginsel	principle	1.12	5.75	insekt	insect	6.57	7.00
groot	huge	1.38	6.25	interesse	interest	4.75	5.88
groot	tall	1.62	5.75	invloed	influence	5.00	6.75
grootte	size	1.12	6.88	inwoner	citizen	1.25	5.62
grot	cave	1.12	7.00	inwoner	inhabitant	2.25	6.75
grots	giant	2.88	5.38	inwoner	resident	1.25	6.50
gunst	favour	1.25	6.88	inzicht	insight	5.75	6.00
haai	shark	1.62	7.00	inzicht	understanding	1.12	6.12
haar	hair	5.25	6.88	jaar	year	5.88	7.00

APPENDIX B (Continued)

Dutch Word	English Word	Form Similarity	Semantic Similarity
jammer	pity	1.12	6.62
japon	dress	1.25	6.00
japon	gown	1.25	6.62
jas	coat	1.25	6.50
jas	jacket	3.25	5.25
jeugd	youth	5.00	6.62
jongen	boy	1.38	6.88
jongen	lad	1.00	6.25
jurk	dress	1.25	7.00
jurk	gown	1.25	6.25
kaak	jaw	1.25	6.88
kaars	candle	3.62	7.00
kaart	map	1.12	6.88
kaartje	card	4.50	6.50
kaartje	postcard	2.62	6.38
kaartje	tag	1.25	5.25
kaartje	ticket	2.12	6.75
kado	gift	1.50	6.50
kalf	calf	5.75	7.00
kamer	room	1.12	6.88
kans	chance	5.50	6.88
kans	opportunity	1.62	6.38
kantoor	office	1.25	6.88
kapitein	captain	5.88	6.62
kast	chest	4.00	4.12
kat	cat	6.25	7.00
katoen	cotton	5.00	7.00
keel	throat	1.75	6.50
kerel	chap	1.38	6.50
kerel	dude	1.00	6.38
kerel	fellow	1.62	6.12
kerel	guy	1.12	6.50
kerel	lad	1.25	6.50
keten	chain	1.88	6.75
keten	string	1.12	4.75
ketting	chain	1.25	6.88
keuken	kitchen	4.12	7.00
keus	choice	2.00	6.75
keuze	choice	3.62	6.75
kijken	watch	1.25	6.75
kijkt	watch	1.25	5.25
kikker	frog	1.62	6.88
kin	chin	6.00	6.75
klacht	complaint	1.88	7.00
kleur	colour	3.50	7.00
klimaat	climate	5.75	6.88
klok	clock	5.75	7.00
kloof	canyon	2.12	5.88
kloof	gap	1.12	5.75
kloof	gorge	1.71	5.43
knoflook	garlic	1.38	6.62
knul	lad	1.00	6.50
koe	cow	3.75	7.00
kofferbak	trunk	1.12	6.88
koffie	coffee	6.00	7.00
kogel	bullet	1.75	7.00
komst	arrival	1.12	6.00
konijn	rabbit	1.25	6.88
koning	king	4.88	7.00
koningin	queen	3.38	6.75
kooi	cage	3.50	6.88
koorts	fever	1.12	7.00
koren	corn	4.88	2.75
kostuum	costume	5.38	6.38
kostuum	suit	1.25	6.38
kraan	faucet	1.00	6.00
kraan	tap	1.12	6.75
kracht	force	1.38	6.25
kracht	power	1.25	6.62
kracht	strength	1.50	6.62
krant	newspaper	1.12	7.00
krant	paper	1.25	6.38
kritiek	criticism	4.25	6.62
kroon	crown	5.12	6.88
kruid	herb	1.25	6.12
kuit	calf	2.25	3.62
kunst	art	1.25	6.88
kussen	pillow	1.12	7.00
kwaad	anger	1.25	6.25
kwaadheid	anger	1.50	6.50
kwaliteit	quality	5.38	7.00
laars	boot	1.38	6.50
lafaard	coward	1.38	6.75
laken	sheet	1.38	6.62
lam	lamb	6.14	6.57
lammetje	lamb	4.25	7.00
land	country	1.12	6.62
land	land	6.75	6.25
landkaart	map	1.50	6.75
lang	long	5.25	6.50
lang	tall	1.75	6.50
lawaai	noise	1.12	6.50
leeftijd	age	1.12	6.88
leen	loan	5.25	5.88
leeuw	lion	2.88	7.00
leger	army	1.25	6.88
leider	leader	6.00	6.88
lening	loan	4.29	7.00
lens	lens	7.00	6.88
lepel	spoon	1.12	7.00
les	lesson	4.12	6.75
leugen	lie	2.12	6.88
leuk	funny	1.38	6.00
leuk	nice	1.25	6.50
lichaam	body	1.00	6.75
lied	song	1.25	6.88
liedje	song	1.57	6.57
lijf	body	1.38	6.38
litteken	scar	1.12	6.88
loon	payment	1.00	6.38
loon	salary	1.25	6.50
loon	wage	1.00	6.38
loon	wages	1.12	6.50
lot	fate	1.62	6.62
lucht	air	1.12	6.75
lucht	sky	1.00	7.00
maag	stomach	1.25	6.88
maan	moon	5.38	7.00
maat	size	1.12	6.25
macht	power	1.12	6.50
mais	corn	1.12	6.88
manier	manner	5.50	6.50
manier	mode	2.12	5.62
manier	way	1.25	6.25
masker	mask	5.88	7.00
mast	mast	6.88	5.50
medaille	medal	4.62	6.75
medelijden	mercy	2.12	6.50
medelijden	pity	1.12	6.50
meelij	pity	1.12	6.62
meerderheid	majority	3.62	6.88
meester	master	6.00	6.50
meid	chick	1.25	4.88
meid	girl	1.12	5.75
meid	maid	6.12	3.88
meisje	girl	1.25	7.00

APPENDIX B (Continued)

Dutch Word	English Word	Form Similarity	Semantic Similarity
meneer	gentleman	2.00	6.25
menigte	crowd	1.12	6.62
mening	meaning	5.88	3.62
mening	opinion	1.38	6.88
mensen	people	1.25	6.62
mes	knife	1.12	7.00
metaal	metal	5.88	6.88
methode	method	5.88	7.00
middelpunt	middle	4.88	5.75
middelste	middle	4.50	6.25
midden	middle	5.00	6.62
mier	ant	1.38	7.00
mijl	mile	6.25	6.62
misbruik	abuse	1.29	6.29
misdaad	crime	1.12	6.62
misdadiger	crook	1.12	5.88
mislukking	failure	1.38	6.38
mist	fog	1.62	6.50
mist	mist	7.00	7.00
modder	mud	4.25	7.00
mode	fashion	1.12	7.00
mode	mode	6.88	4.88
modus	mode	5.43	5.57
moeder	mother	5.88	7.00
moedig	bold	1.12	5.12
mogelijkheid	chance	1.50	5.62
mogelijkheid	opportunity	1.75	6.25
mogelijkheid	possibility	2.00	6.88
molen	mill	3.62	7.00
molen	windmill	1.88	5.62
mond	mouth	4.25	6.88
monnik	monk	5.50	6.88
mooi	fair	1.38	4.88
moordenaar	murderer	4.25	6.88
motief	motive	5.88	6.25
mouw	sleeve	1.12	6.88
muildier	mule	4.71	6.43
muis	mouse	4.38	7.00
munt	coin	1.12	7.00
muntje	coin	1.12	7.00
muur	wall	1.50	6.88
muziek	music	5.25	7.00
naald	needle	4.25	7.00
naam	name	6.12	6.88
nacht	night	4.75	6.75
nadeel	disadvantage	1.00	6.88
nat	wet	1.88	7.00
natuur	nature	6.00	6.62
nederlaag	defeat	1.12	6.62
neger	negro	5.12	6.38
nek	neck	6.25	7.00
neus	nose	4.38	7.00
nobel	noble	6.25	6.62
nonsens	nonsense	6.43	6.71
noodlot	destiny	1.12	6.00
noodlot	fate	1.25	6.25
noodzaak	necessity	2.00	6.38
noodzaak	need	2.50	6.00
noodzakelijkheid	necessity	2.12	6.75
ochtend	dawn	1.25	5.62
ochtendschemer	dawn	1.12	6.38
offer	offer	6.88	3.25
offer	sacrifice	1.25	6.75
onderbroek	pants	1.00	3.75
onderwijs	education	1.38	6.88
oneven	irregular	1.38	4.25
oneven	odd	1.50	6.12
oneven	unequal	2.12	5.12
oneven	uneven	5.12	6.38
onlust	riot	1.38	4.00
onschuld	innocence	1.50	7.00
ontdekking	discovery	1.75	6.75
ontmoeting	meeting	3.38	6.50
onzin	nonsense	1.62	7.00
oog	eye	4.00	7.00
oom	uncle	2.62	7.00
oor	ear	4.25	7.00
oorlog	war	1.25	7.00
oorzaak	cause	1.38	6.75
opdracht	task	1.25	6.38
openbaar	public	1.12	6.75
opening	gap	1.00	6.00
opinie	opinion	5.62	6.00
opleiding	education	1.50	6.75
oplichter	conman	1.38	6.50
oplichter	crook	1.25	6.12
oplossing	solution	1.75	6.50
opofferen	sacrifice	1.25	6.75
opoffering	sacrifice	1.25	6.50
opstand	riot	1.25	6.12
order	order	6.75	5.38
overeenkomst	treaty	1.00	6.12
overjas	cloak	1.43	5.14
overstroming	flood	1.38	6.62
paal	pole	4.62	5.75
paar	couple	1.00	6.57
paar	pair	5.88	6.50
paard	horse	1.25	7.00
pak	suit	1.14	6.71
paniek	panic	5.50	6.75
papier	sheet	1.12	5.62
paraplu	umbrella	1.25	6.88
parel	pearl	5.50	7.00
passie	passion	4.88	7.00
paus	pope	2.75	6.88
peer	pear	5.88	6.88
pen	pen	7.00	6.75
peper	pepper	6.00	7.00
perzik	peach	3.38	7.00
pijl	arrow	1.12	6.75
pijp	pipe	5.88	6.62
pin	pin	7.00	6.00
piraat	pirate	5.25	6.88
pistool	gun	1.00	6.75
pistool	pistol	6.12	6.75
plaats	town	1.25	5.38
plaatsbewijs	ticket	1.12	5.75
plafond	ceiling	1.12	7.00
plan	idea	1.12	5.12
plan	plan	7.00	7.00
plattegrond	map	1.00	6.75
plezier	joy	1.38	6.50
plicht	duty	1.00	6.50
poging	attempt	1.12	6.75
poging	effort	1.25	6.00
poging	try	1.00	5.88
pokken	pox	5.38	6.62
politie	police	5.38	6.88
pond	pound	6.14	6.00
pool	pole	5.75	4.75
pop	doll	1.75	6.88
pop	puppet	3.62	6.88
post	mail	1.12	7.00
post	post	6.75	3.00
pot	jar	1.75	6.50
pot	pot	7.00	5.00

APPENDIX B (Continued)

Dutch Word	English Word	Form Similarity	Semantic Similarity
potlood	pencil	2.25	6.88
preek	speech	2.75	4.88
prijs	price	5.62	7.00
prijs	prize	5.43	6.86
principe	principle	6.12	6.25
prins	prince	6.25	6.88
publiek	audience	1.00	6.88
publiek	crowd	1.12	6.38
publiek	public	5.50	6.75
publiekelijk	public	4.25	6.75
raadsel	riddle	4.50	6.75
raam	window	1.00	7.00
raar	odd	1.50	6.00
rail	rail	6.88	6.75
rail	rails	5.88	6.50
rauw	raw	4.75	6.38
recent	current	2.25	6.00
rechten	law	1.00	5.75
reden	cause	1.12	6.38
reden	reason	3.00	6.88
regel	line	1.50	6.50
regel	rule	4.75	6.62
regel	sentence	1.62	4.88
regen	rain	3.25	6.88
rel	riot	3.00	6.88
rente	interest	1.62	6.75
respect	respect	7.00	6.62
resultaat	result	5.88	6.75
reuk	scent	1.25	6.50
reuk	smell	1.00	6.88
reus	giant	1.29	6.75
richting	direction	1.75	6.75
ridder	knight	1.50	6.88
riem	belt	1.12	6.88
rietje	straw	1.25	6.38
rijst	rice	5.25	7.00
ritme	rhythm	5.00	6.88
rivier	river	5.88	7.00
roem	fame	1.38	6.62
rok	skirt	1.25	6.75
romp	torso	2.25	6.38
romp	trunk	1.12	5.25
rondje	circle	1.25	5.62
rook	smoke	3.38	6.38
roos	rose	5.88	7.00
rouw	grief	1.38	5.75
ruggegraat	spine	1.29	6.86
rugtas	bag	1.38	5.12
ruig	rough	4.00	6.50
ruiken	smell	1.38	6.50
ruil	exchange	1.12	6.88
ruil	swap	1.00	6.43
ruil	trade	1.12	5.62
ruilen	exchange	1.00	5.86
ruilen	trade	1.25	6.88
ruilhandel	trade	1.00	5.88
ruimte	space	1.00	6.50
ruw	crude	2.00	5.88
ruw	raw	4.38	4.62
ruw	rough	3.88	5.88
ruwheid	cruelty	1.12	4.38
ruzie	fight	1.12	5.88
ruzie	quarrel	1.00	6.62
ruzie	riot	3.00	5.12
saai	boring	1.25	6.75
saai	dull	1.38	6.38
salaris	wage	1.25	6.38
sap	juice	1.12	6.88
saus	sauce	6.25	6.75
schaamte	shame	4.00	6.62
schaap	sheep	6.00	6.88
schaar	scissors	2.25	7.00
schandaal	scandal	5.50	6.75
schatting	estimate	1.38	6.50
schatting	estimation	1.25	6.62
schemer	dusk	1.25	6.38
schemer	twilight	1.50	6.62
schemering	dawn	1.12	6.62
schemering	dusk	1.50	6.75
schemering	twilight	1.25	6.62
schepping	creation	1.25	6.75
schijterd	coward	1.12	6.25
schildpad	turtle	1.25	7.00
schoen	boot	2.12	4.12
schoen	shoe	5.38	7.00
schoonheid	beauty	1.38	6.88
schotel	saucer	2.50	5.88
schoteltje	saucer	2.50	6.38
schouder	shoulder	5.50	6.88
schrijver	author	1.00	7.00
schuld	blame	1.25	6.75
schuld	debt	1.50	6.62
schuld	guilt	2.50	7.00
seizoen	season	5.62	7.00
sheet	sheet	7.00	5.12
simpelheid	simplicity	5.00	6.12
sjaal	scarf	2.88	6.00
sjaal	shawl	5.62	6.50
slaaf	slave	5.12	6.88
slachtoffer	victim	1.00	6.88
slager	butcher	1.38	6.75
slang	snake	2.50	7.00
slim	smart	1.88	6.62
smaak	flavour	1.25	6.75
smaak	taste	1.71	7.00
sneeuw	snow	5.62	7.00
snoer	wire	1.12	6.75
snor	moustache	1.00	7.00
sok	sock	6.25	6.25
spanning	excitement	1.62	6.38
spanning	tension	1.25	6.88
speld	pin	1.25	6.62
spiegel	mirror	1.25	7.00
spijt	regret	1.12	6.88
spoor	rail	1.25	6.75
spraak	speech	3.25	6.62
sprookje	tale	1.25	5.62
staart	tail	1.38	6.88
staat	shape	3.50	4.38
stad	city	1.38	6.75
stad	town	1.00	6.62
stam	trunk	1.25	5.25
stand	mode	1.12	4.75
stank	stench	3.38	6.50
stapel	pile	1.38	6.75
steeg	alley	1.12	7.00
steegje	alley	1.25	7.00
steen	rock	1.12	6.88
steen	stone	5.50	7.00
stem	voice	1.50	6.50
stem	vote	1.25	6.88
sterkte	strength	3.25	6.00
steun	support	1.75	6.88
stier	bull	1.38	6.62
stilte	silence	2.38	7.00
stoel	chair	2.25	6.88

APPENDIX B (Continued)

Dutch Word	English Word	Form Similarity	Semantic Similarity
stoer	tough	1.75	5.88
stof	dust	2.00	6.38
stof	fabric	1.25	6.50
stof	material	1.14	5.75
stok	pole	1.62	5.50
straat	road	1.50	6.38
straat	street	6.12	7.00
strand	beach	1.88	7.00
strijken	iron	1.12	5.88
stro	straw	4.38	6.88
stroming	current	1.25	4.88
stroming	flow	1.25	5.50
stroming	stream	5.38	5.75
stroom	current	1.50	6.12
stroom	flood	2.00	6.38
strootje	straw	2.62	6.62
strot	throat	3.12	6.00
stuk	piece	1.88	6.25
suiker	sugar	4.75	6.88
taai	tough	3.00	6.38
taak	duty	1.14	6.14
taak	job	1.12	6.00
taak	task	4.62	6.88
taal	language	1.00	6.75
taart	cake	1.50	5.75
taart	pie	1.00	6.62
tabak	tobacco	4.75	7.00
tafel	desk	1.38	4.75
tante	aunt	2.12	7.00
tap	tap	6.88	5.75
tas	bag	1.88	6.88
taxi	cab	1.12	6.62
taxi	taxi	6.86	7.00
teken	sign	1.12	6.88
teken	token	5.88	3.88
tekort	lack	1.62	6.38
tekort	shortage	2.25	6.75
testament	will	1.00	6.62
thee	tea	5.50	7.00
ticket	ticket	7.00	6.38
tijd	time	5.12	6.88
tijger	tiger	6.12	7.00
toekomst	future	1.38	6.88
toespraak	speech	2.00	6.50
toevluchtsoord	haven	1.12	6.50
toevluchtsoord	refuge	1.25	6.75
tong	tongue	6.00	6.88
toorts	torch	4.75	6.88
touw	rope	1.50	6.50
traan	tear	3.86	6.86
trein	train	5.88	7.00
triangel	triangle	5.88	4.38
troep	trash	2.75	6.38
trots	pride	1.00	6.75
trots	proud	1.25	7.00
tuin	garden	1.12	7.00
tweeling	twin	2.62	7.00
tweeling	twins	2.88	6.88
uitstel	delay	1.62	6.00
uitverkoop	sale	1.12	6.75
uitwisselen	exchange	1.50	6.75
uitwisseling	exchange	1.00	6.88
vaardigheid	ability	1.50	6.62
vaardigheid	skill	1.00	6.75
vaas	vase	6.12	7.00
vacht	fur	1.62	6.75
vader	dad	1.62	6.25
vader	father	5.62	6.88
varken	pig	1.12	7.00
vastbinden	pin	1.62	5.12
vel	sheet	1.00	5.38
veld	domain	1.12	3.62
veld	field	3.88	6.62
verandering	change	1.25	6.50
verdelen	part	1.00	5.75
verdrag	convention	1.38	6.38
verdrag	treaty	1.50	6.88
verdriet	grief	2.38	6.12
verdriet	pain	1.12	4.25
verdriet	sadness	1.00	6.50
verdriet	sorrow	1.38	6.25
verf	paint	1.38	6.88
vergadering	meeting	1.25	6.00
verhaal	story	1.38	6.88
verhaal	tale	1.38	6.50
verjaardag	birthday	1.38	7.00
verkoop	sale	1.25	6.88
verkoop	sell	1.38	6.25
verkoudheid	flu	1.62	4.62
verlegen	shy	1.38	6.88
verlies	defeat	1.50	6.00
verpleegster	nurse	1.12	7.00
verraad	betrayal	1.50	7.00
verraad	treason	1.00	6.12
verrassing	surprise	1.43	6.88
verschil	difference	1.38	7.00
verslagen	defeat	1.62	6.25
vertraging	delay	1.75	6.62
vertrouwen	faith	1.62	6.62
viezigheid	dirt	1.25	5.88
vijand	enemy	1.25	7.00
vinger	finger	6.38	7.00
viool	violin	5.25	7.00
vlam	flame	4.75	6.75
vlees	meat	1.25	6.50
vleugel	wing	1.62	6.88
vlieg	fly	3.50	6.75
vliegen	fly	3.25	6.62
vliegtuig	airplane	1.00	7.00
vliegtuig	plane	1.38	6.00
vlinder	butterfly	1.25	7.00
vlo	flea	3.38	7.00
vloed	flood	5.00	6.12
vloek	curse	2.00	6.75
vloek	spell	1.50	6.00
voet	foot	5.88	6.75
vogel	bird	1.25	6.75
volgorde	order	2.38	6.00
volk	people	1.12	6.38
voorbeeld	example	1.25	6.88
voordeel	advantage	1.38	6.88
voordracht	speech	1.12	6.12
voorhoofd	brow	1.12	3.75
voorhoofd	forehead	4.12	6.75
voorkeur	favour	2.25	6.62
voorstel	proposal	1.75	6.75
voorstel	proposition	1.25	6.75
voortgang	progress	1.12	6.62
vooruitgang	advancement	1.50	6.25
vooruitgang	progress	1.43	6.57
vordeel	favour	1.62	6.25
vordering	progress	1.75	6.75
vorm	form	5.75	5.25
vorm	shape	1.12	6.62
vos	fox	4.75	7.00
vraag	demand	1.25	5.62

APPENDIX B (Continued)

Dutch Word	English Word	Form Similarity	Semantic Similarity
vraag	question	1.00	6.75
vrede	peace	1.50	7.00
vreemd	odd	1.12	6.00
vreugde	happiness	1.25	6.38
vreugde	joy	1.25	6.75
vriend	chap	1.25	5.25
vriend	friend	5.88	6.75
vrouw	dame	1.25	4.75
vrouw	female	1.50	6.88
vrouw	lady	1.25	5.62
vrouw	woman	2.12	7.00
vuil	dirt	1.14	6.00
vuil	trash	1.25	6.00
vuilnis	trash	1.25	6.38
vuist	fist	4.12	7.00
vuur	fire	3.38	7.00
waard	worth	3.88	6.62
waarde	value	1.38	7.00
waarde	worth	3.62	6.38
waarheid	truth	1.25	6.88
wachter	watch	4.38	5.12
walvis	whale	4.88	7.00
wand	wall	4.57	6.29
wang	cheek	1.25	6.50
wanhoop	despair	1.25	6.50
want	glove	1.00	5.50
warmte	heat	1.25	6.00
warmte	warmth	5.62	6.88
water	water	6.88	7.00
weg	road	1.12	6.88
wegkwijnen	pine	1.88	5.25
wenkbrauw	brow	3.14	5.88
wens	wish	3.12	6.75
wereld	world	4.62	6.88
werkster	maid	1.12	6.88
werktuig	tool	1.00	6.50
werktuig	utensil	1.29	5.62
west	west	7.00	6.88
westen	west	5.62	7.00
wet	law	1.62	6.88
wetenschap	science	1.62	6.50
wiel	wheel	5.38	6.88
wijsheid	knowledge	1.50	6.00
wijsheid	wisdom	3.62	6.62
wil	will	6.50	5.88
wind	wind	7.00	7.00
winkel	shop	1.50	6.62
winkel	store	1.38	6.25
winkelen	shop	1.38	6.25
winter	winter	6.88	7.00
wissel	change	1.38	5.12
wisselgeld	change	1.25	6.00
wisseling	chance	1.12	3.75
woede	anger	1.12	6.14
wolf	wolf	6.00	7.00
wolk	cloud	1.38	7.00
woord	word	6.00	6.86
wortel	carrot	1.62	6.62
wortel	root	2.00	6.88
woud	forest	1.12	6.38
woud	wood	5.25	5.88
woud	woods	4.50	5.50
wraak	revenge	1.25	6.75
wraak	vengeance	1.38	6.62
wreed	crude	1.38	5.38
wreed	cruel	1.12	6.75
wreedheid	cruelty	1.38	6.75
wrok	grudge	1.25	6.62
zaak	case	1.62	6.50
zak	bag	3.62	6.25
zak	pocket	1.12	5.88
zal	will	1.50	6.50
zee	sea	6.00	7.00
zeep	soap	5.62	6.88
zeil	sail	5.38	6.50
zeilen	sail	4.38	6.75
zekerheid	certainty	3.00	6.38
zekerheid	security	2.38	5.88
ziekenhuis	hospital	1.62	6.75
zijde	silk	3.62	6.88
zilver	silver	6.25	6.88
zomer	summer	5.25	7.00
zon	sun	6.00	7.00
zonde	pity	2.12	5.50
zonde	sin	3.25	6.75
zondigen	sin	2.38	6.50
zonsopgang	dawn	1.12	6.62
zonsopkomst	dawn	1.12	6.25
zoon	son	4.50	6.88
zorg	care	1.75	6.38
zorgen	care	1.12	6.50
zout	salt	4.14	7.00
zullen	will	1.12	5.62
zus	sister	2.88	6.75
zuster	nurse	1.88	6.75
zuster	sister	4.75	7.00
zwaar	rough	1.50	6.00
zwaar	tough	1.62	5.62
zwakheid	weakness	1.50	6.75
zwakte	weakness	1.75	6.75
zweet	sweat	5.38	6.88
zweten	sweat	4.88	6.88

Note—For complete norms, see http://www.talkbank.org/norms/tokowicz/.

(Manuscript received November 6, 2001;
revision accepted for publication May 26, 2002.)