Ambiguous words are harder to learn*

TAMAR DEGANI
NATASHA TOKOWICZ
University of Pittsburgh

(Received: 23 April 2009; Revised: 15 October 2009; Accepted: 22 October 2009; First published online 19 January 2010)

Relatively little is known about the role of ambiguity in adult second-language learning. In this study, native English speakers learned Dutch–English translation pairs that either mapped in a one-to-one fashion (unambiguous items) in that a Dutch word uniquely corresponded to one English word, or mapped in a one-to-many fashion (ambiguous items), with two Dutch translations corresponding to a single English word. These two Dutch translations could function as exact synonyms, corresponding to a single meaning, or could correspond to different meanings of an ambiguous English word (e.g., wisselgeld denotes the monetary meaning of the word change, and verandering denotes alteration). Several immediate and delayed tests revealed that such translation ambiguity creates a challenge for learners. Furthermore, words with multiple translations corresponding to the same meaning are more difficult to learn than words with multiple translations corresponding to multiple meanings, suggesting that a one-to-many mapping underlies this ambiguity disadvantage.

Many adults who try to learn a second language (L2) report great difficulty in doing so. A major barrier to their success is the fact that there are often mismappings between the first language (L1) and L2, at a variety of levels (see e.g., Tokowicz, Kroll, de Groot & van Hell, 2002, for the lexical level; Tokowicz & MacWhinney, 2005, for morphosyntax; see also MacWhinney, 1997). Here, we explored the impact of a particular kind of mismapping between L1 and L2 on vocabulary learning. Specifically, we explored TRANSLATION AMBIGUITY, which occurs when a word in one language has more than one translation into another language. Many instances of translation ambiguity arise from two existing kinds of ambiguity within a language. The first source of translation ambiguity is near-synonymy, which occurs when a given meaning can be expressed by more than one word form (e.g., *couch* and *sofa*). The second occurs when a given word form corresponds to more than one meaning (homonymy) or sense (polysemy), as in the case of the English homonym *change*, which can denote both "coins of small denomination" and "the result of alteration".

Although there is a paucity of research on adult learning of ambiguous words, children have long been shown to have difficulty when there is not a one-to-one mapping between words and meanings, as in the case of learning homonyms and near-synonyms. For example, Doherty (2004) examined children's ability to learn a new meaning for a pseudo-homonym. The word *fork*, for instance, was introduced in a story as referring to a different object. Children were significantly less accurate at selecting the intended referent when a pseudo-homonym served as the referent's label compared to when a nonsense word was used. Further, when children's attention was specifically drawn to the fact that a word may refer to two different things, children were still less accurate at labeling the referent of the pseudo-homonym when the object to which the word typically refers (i.e., an actual fork) was present. Children therefore have difficulty in associating a new meaning to a word they already know (see also Mazzocco, 1997).

Traditionally, this difficulty in learning one-to-many mappings has been ascribed to a mutual exclusivity bias, according to which children prefer unique mappings between word form and meaning, and require more evidence to learn words that violate this preference (e.g., Markman & Wachtel, 1988). Obviously, children learn to overcome this difficulty with time (e.g., Mazzocco, 1997), and multilingual children need to learn from the outset that the same object can have more than one label (e.g., Byers-Heinlein & Werker, 2009). Nonetheless, a one-to-many mapping may continue to pose difficulty in learning even for adults (Golinkoff, Hirsh-Pasek, Bailey & Wenger, 1992), perhaps by means of a different mechanism, which will be considered in the Discussion section, below.

The current study focuses on ambiguity that crosses the language boundary, in that the ambiguity exists in the mapping from a word in one language (English) to more than one word in the other language (Dutch). Such translation ambiguity is relatively prevalent in several cross-language pairs, including English and Dutch (Tokowicz et al., 2002; see also Prior, MacWhinney & Kroll, 2007; and Tokowicz & Kroll, 2007,

* This project was supported by NSF-BCS 0745372 and a Language Learning Grant awarded to NT. We thank the members of the PLUM Lab and Kevin Jarbo, Adiam Mekonen, and Angela Sperl for research assistance, and Charles A. Perfetti and two anonymous reviewers for their comments on an earlier version of this manuscript. A version of this research was presented at the 49th Annual Meeting of the Psychonomic Society, in November 2008.

Address for correspondence:
Natasha Tokowicz, Learning Research & Development Center, 3939 O'Hara St., Room 634, University of Pittsburgh, Pittsburgh, PA 15260, USA
tokowicz@pitt.edu

Table 1. *Example stimuli and definitions by condition.*

Condition	English word	Definition(s)	Dutch translation(s)
Form-ambiguous	sky	1. the region of the clouds or the upper air	1. lucht 2. hemel
Meaning-ambiguous	change	1. the result of alteration or modification 2. coins of small denomination	1. verandering 2. wisselgeld
Unambiguous	arrow	1. a mark with a pointed end used to indicate a direction or relation	1. pijl

for English/Spanish). Semantic (meaning) ambiguity and near-synonymy contribute to such cross-language ambiguity. Specifically, when a word has more than one meaning (e.g., *change*) it is rarely the case that one translation in the other language also happens to capture these same two meanings (Frenck-Mestre & Prince, 1997), which are thought to have been accidentally joined (e.g., Rodd, Gaskell & Marslen-Wilson, 2002). Thus, a given word in English (*change*) corresponds to two translations in Dutch: *wisselgeld* refers to the monetary meaning of *change*, whereas *verandering* denotes its alteration meaning. Because the English word is ambiguous in its meaning, we refer to this type of ambiguity as MEANING AMBIGUITY. Conversely, the English word *sky* can be translated into one of the two Dutch near-synonyms *lucht* and *hemel*. This near-synonymy in Dutch creates cross-language ambiguity that we refer to as FORM AMBIGUITY, because when translating the English word into Dutch there is relatively no ambiguity in its meaning but rather two word forms are available, placing the ambiguity at the word-form level (see Table 1 for example stimuli).

Translation ambiguity has been shown to affect bilingual language processing. In particular, Tokowicz et al. (2002) found that translation-unambiguous pairs, which map uniquely to one another, were rated as more similar in meaning than translation-ambiguous pairs, which are made up of words that can be translated in more than one way. For example, *frog* is only translated into Dutch as *kikker*, and vice versa, and this pair received a higher meaning similarity rating than both *change–wisselgeld* and *change–verandering*. Moreover, Tokowicz and Kroll (2007) showed that translation-ambiguous pairs were produced more slowly and less accurately in a translation production task by relatively proficient Spanish–English bilinguals, likely reflecting active competition between the two possible output options (see also Tokowicz, Michael & Smith, 2007; Tokowicz, Prior & Kroll, 2009). This slowing of translation production for words with more than one translation became even more pronounced with increased L2 proficiency (Tokowicz, 2005); presumably, this proficiency difference is due to increased word knowledge that leads to stronger competition between alternatives.

Recently, Boada, Sánchez-Casas, García-Albea, Gaviln and Ferr (2009) examined processing of translation-ambiguous words in a translation recognition task, in which participants indicated whether pairs of words were translations of each other. Translation-ambiguous pairs were correctly recognized as translations more slowly than translation-unambiguous pairs, demonstrating that translation ambiguity affects recognition as well as production (see also Sánchez-Casas, Buratti & Igoa, 1992).

Despite the prevalence of ambiguity within and across languages, and its effects on the way monolinguals and bilinguals process language (e.g., Elston-Güttler & Friederici, 2005; for a review, see Degani & Tokowicz, in press), and on how children learn vocabulary, very little attention has been paid to the effect of ambiguity on adult language learning. In one study that examined foreign vocabulary learning in adolescents, Bogaards (2001) in effect manipulated the mapping between words and meanings. He contrasted learning of completely new words with learning new meanings for already-known word forms. Specifically, native Dutch speakers learned idioms made up of known French words, learned new meanings for already-known French words, or learned entirely new words. Using accuracy on a multiple-choice translation recognition task and a translation production task, Bogaards found that learning a new meaning for an already-known form is easier than learning a completely new word.

The results of the Bogaards (2001) study suggest an advantage, rather than a disadvantage, for words with more than one meaning over unambiguous words. Such an advantage was also observed in a study with children that demonstrated an accuracy advantage for words with one form and two meanings over completely novel words in a picture naming task (Storkel & Maekawa, 2005). The authors took their results to suggest that the reduced amount of information that needs to be learned for ambiguous words (because the word form

is already known) is what provides words with two meanings an advantage over unambiguous novel words. The discrepancy between studies that demonstrate a disadvantage for ambiguous words (e.g., Doherty, 2004) and studies that find an ambiguity advantage (e.g., Bogaards, 2001; Storkel & Maekawa, 2005) is explained by differences in the types of representations that the tasks access.

Notably, the studies that found an ambiguity advantage in fact compared learning of novel words, which requires creation of an additional form representation, to learning of new meanings for already-known words, for which no new form representations must be created. Thus, because these previous studies were not focused on comparing ambiguous and unambiguous word learning, ambiguous and unambiguous items differed in the number of representations that needed to be created by the learner.

In the current study, we controlled for the number of representations that needed to be created. Specifically, we taught words in Dutch that did or did not share a common English translation. Therefore, by virtue of being native English speakers, the participants in the current study were familiar with the English ambiguous items and the meanings they encompass. Thus, for unambiguous, form-ambiguous, and meaning-ambiguous items, the English words and the meanings were already known, and only the Dutch translation(s) had to be learned. Thus, the current study measures more directly how one-to-many mappings, PER SE, rather than the amount of information learned, affect learning.

The primary factor of interest in the present study is translation ambiguity. We compare the cases in which one word form in English corresponds to one word form in Dutch (a one-to-one mapping) to the cases in which one word form in English corresponds to two word forms in Dutch (a one-to-many mapping). Therefore, if learning is affected by the MAPPING between word forms in the two languages, and learning a one-to-many mapping is more difficult, then ambiguous words should show a disadvantage relative to unambiguous words.

We further examined whether the ambiguity of the source word (i.e., English) modulates the difficulty with which translation-ambiguous items are learned. Therefore, we included both unambiguous English words (e.g., *sky*) with two Dutch words corresponding to the same meaning (form-ambiguous), and ambiguous English words (e.g., *change*) with Dutch translations for each of their meanings (meaning-ambiguous). These two types of translation ambiguities may have different consequences for word learning. In particular, if one assumes that a meaning-ambiguous word has one lexical entry but two semantic entries (one for each meaning; e.g., Klein & Murphy, 2001), then when learning to map a Dutch translation to each of these meanings, the learner can avoid the one-to-many mapping problem. This is especially true if a direct link can be created from the Dutch (L2) word to the semantic/conceptual representation. Although this typically may not be the case in non-immersion learning (see e.g., Kroll & Stewart, 1994), we always provided the definition along with the English translation during training, which may have allowed the learners to rely less on the English word forms and more on meaning. For the form-ambiguous items, in contrast, only one semantic representation should exist, by virtue of the one definition provided, and therefore a one-to-many mapping would have to be established. Under this framework, we would predict an advantage for meaning-ambiguous items over form-ambiguous items.

Alternatively, it may be the case that despite knowing the two different meanings of semantically-ambiguous words in English, creating a mapping between English and Dutch requires some additional meaning differentiation. If so, learning meaning-ambiguous items may be as problematic as learning form-ambiguous items. The idea that it is difficult to differentiate two meanings that correspond to a shared label is supported by a recent study showing that bilinguals consider two words that share a translation in the other language they know to be more similar in meaning than two words that do not share a label (Degani, Prior & Tokowicz, in press; see also Jiang, 2002, 2004). This may suggest that the two meanings of an ambiguous word are closer in semantic space by virtue of their shared label, and the need to associate each meaning with a separate Dutch translation may require the learner to pull the two meanings apart.

It is also possible that ambiguous items are harder to learn because each of their alternatives is encountered less often than unambiguous items. Indeed, this is likely the situation in actual language learning. We wanted to examine, however, whether translation-ambiguous items are more difficult to learn, above and beyond the expected decrement due to less-frequent exposure to each pairing of translations. Therefore, we manipulated the number of repetitions of the ambiguous items, such that ambiguous items could be presented half as often or exactly as often as the unambiguous items. This allows us to examine the effects of translation ambiguity both controlled for and confounded with frequency-of-exposure effects. Equating the number of presentations eliminates the possibility that less-frequent encounters alone underlie the predicted ambiguity disadvantage.

Because some tasks may be more sensitive than others at revealing an ambiguity effect in general, and an ambiguity source effect in particular (see Tokowicz et al., 2009), several tests were employed in the current study. These include a Dutch-to-English translation production task, a translation recognition task, and an English-to-Dutch translation production task. The latter task was

https://doi.org/10.1017/S1366728909990411 Published online by Cambridge University Press

included because translation from L1 to L2 is sometimes more sensitive to meaning effects than translation from L2 to L1 (see e.g., Kroll & Stewart, 1994). In addition to the variety of the tests employed, we also sampled participants' performance at three time points to obtain a more accurate picture of how translation ambiguity affects performance at these very early stages of learning, and to allow us to determine how ambiguity affects performance over time. Because of the potential for frustration due to task difficulty, the English-to-Dutch (L1-to-L2) translation task was administered only during the last session.

To summarize, we expect translation-ambiguous items to have a learning disadvantage as compared to translation-unambiguous items overall. With respect to the source of ambiguity, the various mechanisms we have described make different predictions about the relative difficulty of learning form- and meaning-ambiguous words. If the problem with ambiguous words is that they require a one-to-many mapping, then less difficulty is predicted when learners can avoid a one-to-many mapping. In particular, we would expect that the meaning-ambiguous words would have a learning advantage over form-ambiguous words if learners can directly map each Dutch translation to its corresponding meaning. If, however, meaning-ambiguous words require some meaning differentiation, we would expect them to be almost as problematic as form-ambiguous words, and for unambiguous words to have a learning advantage over both types of ambiguous words.

Method

Participants

Forty-seven right-handed native English speakers (27 males; mean age 18.6 years) with no prior knowledge of Dutch took part in this experiment. They received credit toward an Introductory Psychology requirement and were paid for their participation in the second and/or third sessions. At the end of the first session, the participants completed a language history questionnaire (Tokowicz, Michael & Kroll, 2004; see Table 2). Data from 11 additional participants were excluded because they were native speakers of a language other than English (eight participants) or due to technical problems during training (three participants). Of the final set of 47 participants, 37 returned for a second session two days after the first session, and 17 returned for a third session approximately two and a half weeks after the second.

Design

A two TRANSLATION AMBIGUITY (ambiguous vs. unambiguous) within-participants design was used. For ambiguous items, a two AMBIGUITY SOURCE (form vs. meaning) by two REPETITION LEVEL (2 vs. 4) within-participants design was used.

Table 2. *Language history questionnaire data.*

Measure	Average (SD)
Age (years)	18.66 (1.18)
Age began L2 (years)	12.58 (2.59)
Time studied L2 (years)	4.59 (2.52)
L1 reading ability	9.49 (0.83)
L2 reading ability	4.09 (2.00)
L1 writing ability	9.28 (1.07)
L2 writing ability	3.57 (2.06)
L1 conversation ability	9.70 (0.69)
L2 conversation ability	3.45 (2.21)
L1 speech comprehension ability	9.74 (0.57)
L2 speech comprehension ability	3.82 (2.22)

Note. Reading, writing, conversational, and speech comprehension abilities were rated on a 10-point scale on which 1 indicated the lowest level of ability and 10 indicated the highest level of ability. Standard deviations (SDs) are shown in parentheses. L2 varied across participants but was never Dutch or German.

Stimuli

The stimuli were 48 English words presented with their Dutch translations; the words were selected based on translation norms for Dutch–English translation pairs (Tokowicz et al., 2002), such that half of the English words had a single translation in Dutch, and half had two translations in Dutch. Twelve of the multiple-translation items were meaning-ambiguous, and twelve were form-ambiguous (see Table 1 for examples). Because half of the English words had two translations in Dutch, the stimuli included a total of 72 Dutch words. See the Appendix for all items and their definitions.

The stimuli were selected such that single-translation items were matched to the meaning-ambiguous items and to the form-ambiguous items on English length, English log frequency, and mean reaction time and accuracy to name the English word (all Fs < 1, except mean naming reaction time $p > .26$; taken from Elexicon, Balota et al., 2002), and on familiarity, concreteness, imageability, and age of acquisition of the English words (Fs < 1, from the MRC database, Wilson, 1988; mean characteristics are given in Table 3). For multiple translation items, an average of the length of the two Dutch translations was computed. Single-translation items did not differ significantly from multiple-translation items in Dutch length ($F < 1$).

https://doi.org/10.1017/S1366728909990411 Published online by Cambridge University Press

Table 3. *Properties of stimuli by condition.*

	Condition		
	Form-ambiguous	Meaning-ambiguous	Unambiguous
English length (number of letters)	5.75 (2.38)	5.75 (1.55)	5.75 (1.75)
English log HAL frequency	6.88 (1.64)	6.42 (1.87)	6.54 (2.25)
Familiarity rating	517.67 (165.66)	480.33 (226.50)	534.04 (123.90)
Concreteness rating	399.92 (177.33)	405.75 (216.59)	435.00 (170.31)
Imageability rating	446.17 (180.68)	415.33 (211.74)	471.92 (170.55)
Age of acquisition rating	234.25 (195.41)	204.25 (159.47)	178.71 (180.92)
Mean naming reaction time	622.17 (38.91)	603.93 (40.13)	599.11 (40.11)
Mean naming accuracy	1.00 (0.01)	0.99 (0.02)	1.00 (0.01)
Average Dutch length (number of letters)	7.17 (2.25)	5.75 (2.26)	6.54 (2.25)
Average form similarity rating	2.32 (1.47)	2.22 (1.31)	2.79 (1.97)

Note. Stimuli in the three conditions match on all dimensions. English log frequency, and mean reaction time and accuracy to name the English word are from Elexicon (Balota et al., 2002), and familiarity, concreteness, imageability, and age of acquisition of the English words are from the MRC database (Wilson, 1988). Form similarity ratings are from Tokowicz et al. (2002). Standard deviations (SDs) are shown in parentheses.

Moreover, Tokowicz et al.'s (2002) norms include form similarity ratings for each English–Dutch pair. This measure indicates how 'cognate-like' the pair is; specifically, Dutch–English bilinguals rated the similarity of the pair in spelling and sound on a seven-point scale. Single-translation items did not differ significantly from meaning-ambiguous or form-ambiguous items in their form similarity, which was averaged across the two translations of multiple-translation words ($F < 1$).

An English definition was created for each of the Dutch words by consulting WordNet (WordNet 3.0, Princeton University, 2006) and Dictionary.com (Electronic commerce, 2007) for each of the English items. To keep the definitions short and concise, slight modifications were made on occasion to those provided by the online tools. A separate definition corresponded to each of the meanings (and Dutch translations) of the meaning-ambiguous items. Although the form-ambiguous items may have corresponded to slightly different senses of the English word for proficient Dutch users, we presented both Dutch translations with a single definition, to create a strong one-to-many mapping; we return to this issue in the Discussion section. Single-translation items were also presented with only one definition.

Three training versions were counterbalanced across participants. Each version included 16 multiple-translation English words (eight form-ambiguous and eight meaning-ambiguous), presented with both Dutch translations. The remaining eight multiple-translation items (four form-ambiguous and four meaning-ambiguous) were presented with only one of their Dutch translations (and one definition). These served as fillers to reduce the number of multiple-translation items in each training set. In addition, 24 single-translation items were presented with only one Dutch translation and definition.

Each version therefore included 64 Dutch words to be learned. The number of repetitions of each word was manipulated as well, such that in each training cycle the single-translation items were repeated four times, but of the multiple-translation items presented with both definitions in that version, 16 were repeated four times and 16 were repeated twice. Thus, the corresponding English items appeared eight and four times, respectively, half of the time with each translation. Two variants of each version were created, such that the multiple-translation items that were repeated four times in one variant were repeated twice in the other.

In the Dutch-to-English translation production task in all sessions, all 64 Dutch words trained in that version of the experiment were presented. In the English-to-Dutch translation production test, all 48 English words were presented. The multiple-translation English items that were trained with both Dutch translations in a given participant's corresponding training were presented twice during the test, for a total of 64 trials.

Four different variants of each training version were created so that the order of testing multiple-translation pairs in the translation recognition test was counterbalanced across sessions. Furthermore, pairings were manipulated such that each Dutch word appeared with its correct English translation half of the time; incorrect pairings were made with other items in the "no"

https://doi.org/10.1017/S1366728909990411 Published online by Cambridge University Press

list for that version. Different variants of this task were administered in the second and third sessions.

Procedure

The experiment consisted of three separate sessions, taking place two days and two and a half weeks apart, respectively. During the first session, participants completed two training cycles in which each Dutch word was presented with its English translation and a definition of its meaning. No additional context was provided. On each trial, a fixation cross appeared in the center of the screen until the participant initiated the beginning of the trial by pressing the space bar. The fixation cross was then replaced by a blank screen for 100 ms followed by the Dutch word with its English translation and definition for eight seconds (following Lotto & de Groot, 1998). Participants were instructed to try to learn the Dutch words and their meanings. After two cycles of 208 training trials, each presented in a random order, participants were tested in a Dutch-to-English translation production task, in which they were presented with the Dutch (L2) word and were asked to say its English (L1) translation out loud as quickly and accurately as possible. Following these 64 trials of L2-to-L1 production, participants completed the language history questionnaire.

Two days later, participants returned to the lab for a second training and testing session. In this second session, participants were first tested in the Dutch-to-English translation production task. They then completed one training cycle and were next tested in a translation recognition task. In this task, 48 Dutch–English word pairs were presented in randomized order. On each trial, participants were presented with Dutch–English word pairs, and were asked to indicate whether the English word was a correct translation of the Dutch word by pressing the Yes button with their right index finger or the No button with their left index finger.

Approximately two and a half weeks later ($M = 17$ days, range 14–21), participants returned for a longer-term retention testing session. This third session started with a Dutch-to-English translation production test (L2-to-L1, the same as that administered during the first and second sessions), followed by a translation recognition task (another variant of the task administered in the second session). Participants next completed a working memory span task (the operations-word task, Turner & Engle, 1989; see also Tokowicz et al., 2004) as a delay,[1] before completing an English-to-Dutch translation production task, in which participants orally provided the Dutch (L2) translations of 64 English (L1) words presented in random order.

To summarize, participants completed a total of three training cycles with 208 trials each, and were tested three times in an L2-to-L1 translation production task, twice in a translation recognition task, and once in an L1-to-L2 translation production task. Single-translation (unambiguous) items were trained four times in each cycle, for a total of 12 times in the study. Half of the form- and meaning-ambiguous items were trained twice in each cycle (for a total of six in the study), and the other half were trained four times in each cycle (for a total of 12 in the study).

Results

Data by participants were analyzed using repeated measures Analyses of Variance (ANOVAs) with translation ambiguity, ambiguity source, and repetition level as within-participants variables, and are reported as F_1. Data by items were analyzed using repeated measures ANOVAs with translation ambiguity, ambiguity source, and repetition level as between-items variables, and are reported as F_2. In all the analyses that follow, an alpha level of .05 was used, and in cases in which the assumption of sphericity was violated, the *F*s, *MSE*s, and *p*s reported correspond to the significance test with the Greenhouse-Geisser correction applied.

In the analyses that follow, we first report the comparison of unambiguous items to ambiguous items for which the English word was presented four times in each training cycle. It follows that for the ambiguous items, each Dutch translation was presented only twice (half as often as the Dutch unambiguous items). We then report the analyses when keeping the number of presentations of the DUTCH items at four per cycle, thus equating participants' familiarity with ambiguous and unambiguous Dutch items. These analyses are followed by a more detailed analysis examining the role of ambiguity source, in which we specifically compare form-ambiguous, meaning-ambiguous, and unambiguous items that were all presented four times in each training cycle. An overview of the results can be found in Table 4.

Immediate Dutch-to-English production test: First session

Accuracy

As predicted, ambiguous items were produced less accurately ($M = .60$) than unambiguous items ($M = .75$),

[1] The operations-word task was meant to serve as an individual difference measure of working memory as well as a delay task. Because of a technical problem during task administration, we were unfortunately unable to use these data as an individual difference measure; therefore the task served only as a delay before the English-to-Dutch production test.

https://doi.org/10.1017/S1366728909990411 Published online by Cambridge University Press

Table 4. *Overview of the results.*

Test and session (n)	Condition		
	Form-ambiguous	Meaning-ambiguous	Unambiguous
	Accuracy as % correct (SD)		
Dutch-to-English production 1 (43)	0.69 (0.27) $_{1, a}$	0.67 (0.26) $_{1, a}$	0.75 (0.19) $_{2, a}$
Dutch-to-English production 2 (33)	0.50 (0.24) $_{1, a}$	0.57 (0.29) $_{1, a}$	0.64 (0.19) $_{2, b}$
Dutch-to-English production 3 (17)	0.47 (0.25) $_{1, a}$	0.49 (0.24) $_{1, a}$	0.53 (0.19) $_{1, a}$
Translation recognition 2 (37)	0.95 (0.10) $_{1, a}$	0.94 (0.11) $_{1, a}$	0.97 (0.04) $_{2, a}$
Translation recognition 3 (17)	0.81 (0.23) $_{1, a}$	0.88 (0.13) $_{1, a}$	0.91 (0.07) $_{2, a}$
English-to-Dutch production 3 (17)	0.41 (0.27) $_{1, a}$	0.48 (0.28) $_{1, a}$	0.49 (0.20) $_{1, a}$
	Reaction times in milliseconds (SD)		
Dutch-to-English production 1 (44)	1729 (567) $_{1, a}$	1638 (563) $_{1, a, c}$	1545 (426) $_{2, b, c}$
Dutch-to-English production 2 (32)	1668 (555) $_{1, a}$	1662 (543) $_{1, a}$	1629 (334) $_{1, a}$
Dutch-to-English production 3 (16)	1697 (399) $_{1, a}$	1626 (362) $_{1, a}$	1661 (237) $_{1, a}$
Translation recognition 2 (35)	1534 (446) $_{1, a}$	1304 (327) $_{1, b}$	1242 (236) $_{2, b}$
Translation recognition 3 (16)	1465 (549) $_{1, a}$	1417 (621) $_{1, a}$	1398 (324) $_{1, a}$
English-to-Dutch production 3 (15)	2100 (519) $_{1, a}$	1656 (525) $_{1, b}$	1895 (418) $_{1, b}$

Note. Data are shown as a function of the type of item (form-ambiguous, meaning-ambiguous, or unambiguous) in the equated repetition condition. Means in the same row that do not share a numeric subscript differ at the $p < .05$ level, based on F_1 analyses comparing ambiguous to unambiguous items. Means in the same row that do not share an alphabetic subscript differ at the $p < .05$ level based on *t*-tests between the conditions when the ambiguity source effect was significant by participants. Standard deviations (SDs) are shown in parentheses.

$F_1(1,43) = 56.12$, $MSE = .01$; $p < .01$, $F_2(1,70) = 8.96$, $MSE = .04$, $p < .01$. Furthermore, even when the Dutch ambiguous items were trained just as often as unambiguous items, translation ambiguity led to a learning disadvantage; ambiguous items were still produced significantly less accurately ($M = .68$) than unambiguous items ($M = .75$) by participants, $F_1(1,43) = 8.45$, $MSE = .01$, $p < .01$, and marginally so by items, $F_2(1,70) = 3.02$, $MSE = .03$, $p = .09$. The effect of ambiguity source on accuracy was not significant, $F_1(2,84) = 1.21$, $MSE = .08$, $p > .10$; $F_2(2,69) = 1.54$, $MSE = .03$, $p > .10$.

Reaction times

Based on the distribution of the data, reaction times shorter than 300 ms or longer than 7000 ms, or that fell more than 2.5 standard deviations from each participant's mean reaction time for correct trials were removed from the analyses. These procedures resulted in the removal of 4.9% of the data. Analyses were performed on correct trials only.

The reaction time results from the first session are similar to the accuracy data. Specifically, ambiguous items were produced more slowly ($M = 1795$) than unambiguous items ($M = 1545$), $F_1(1,43) = 23.52$, $MSE = 58310$, $p < .001$; $F_2(1,70) = 6.01$, $MSE = 130482$, $p < .05$. Importantly, as with the accuracy data, equating the number of repetitions of the Dutch ambiguous items during training did not alleviate their disadvantage as compared to unambiguous items; ambiguous items were still produced significantly more slowly ($M = 1685$) than unambiguous items ($M = 1545$), in the analysis by participants, $F_1(1,43) = 6.66$, $MSE = 65273$, $p < .05$, and marginally more slowly in the analysis by items, $F_2(1,70) = 3.01$, $MSE = 128161$, $p = .09$.

A more detailed analysis examining the effect of ambiguity source (meaning vs. form) as compared to unambiguous items revealed a significant effect on learning in the analysis by participants, $F_1(2,86) = 3.39$, $MSE = 109881$, $p < .05$; $F_2(2,69) = 1.69$, $MSE = 129276$, $p > .10$ (see Figure 1). Planned comparisons revealed that only form-ambiguous items were produced more slowly ($M = 1729$) than unambiguous items ($M = 1545$), $t(43) = 2.69$, $p < .05$. Meaning-ambiguous items ($M = 1638$) did not differ significantly from form-ambiguous items, $t(43) = 1.22$, $p > .10$, or from unambiguous items, $t(43) = 1.35$, $p > .10$.

To summarize the production data from the first session, translation-ambiguous items were harder to learn, as measured by the accuracy and reaction time data in an immediate production task. Furthermore, increasing participants' familiarity with ambiguous items by repeating them just as often as unambiguous items

https://doi.org/10.1017/S1366728909990411 Published online by Cambridge University Press

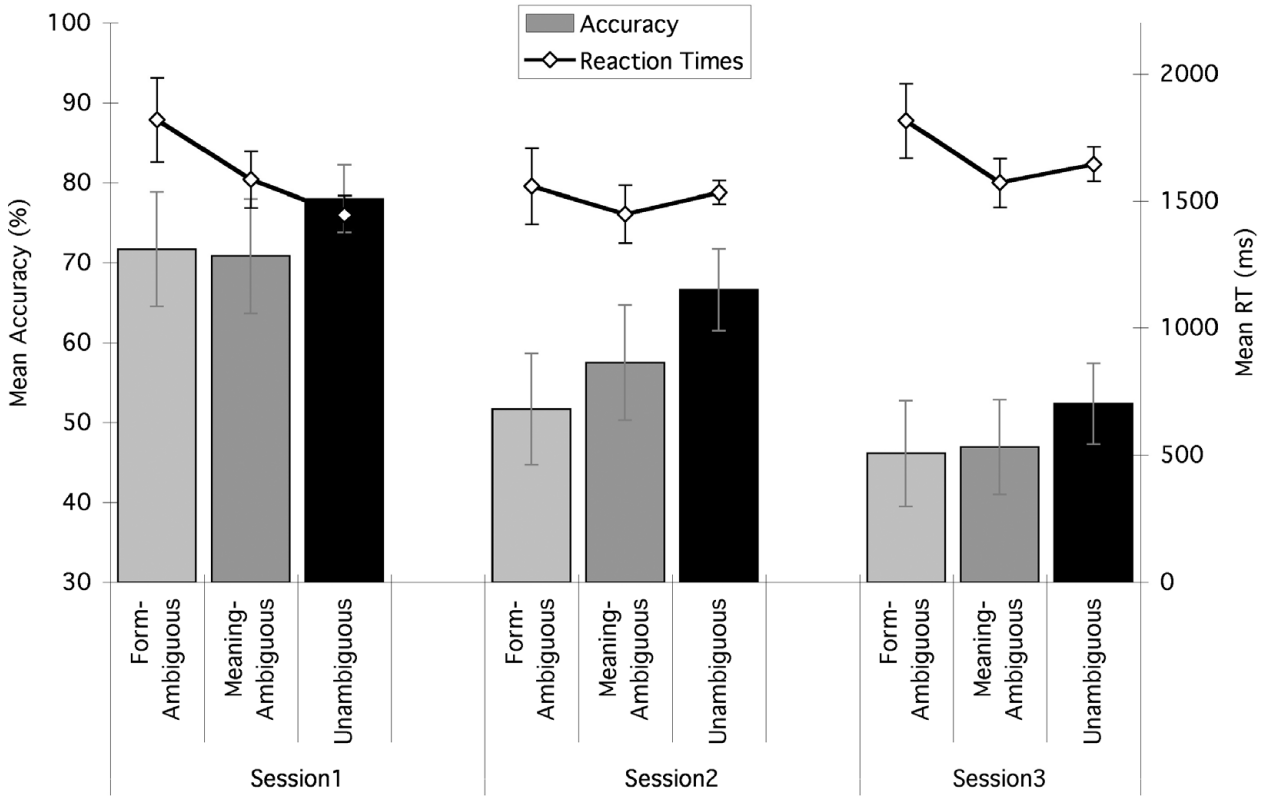

Figure 1. Accuracy and reaction time data from the Dutch-to-English translation production test in the first, second, and third sessions. Data are shown as a function of the type of item (form-ambiguous, meaning-ambiguous, or unambiguous) in the equated repetition condition.

during training did not alleviate this disadvantage. Further, form-ambiguous, but not meaning-ambiguous, items were produced significantly more slowly than unambiguous items.[2]

[2] Because of the long protocol of the study, some participants dropped out before the final session. However, dropout was not balanced with respect to experimental version. To maximize power, we included the maximum number of participants available in each test. By doing so, we were unfortunately unable to maintain the full counterbalancing of the stimuli across the different tests and sessions or to include the same participants in all tests. To examine whether this influenced the pattern of results, we also analyzed the data using experimental version and dropout as factors. In all of the reported analyses, experimental version did not change the interpretation of significant results. Dropout effects were examined by including a between-participant factor of returning to session two and returning to session three as appropriate. In most of the analyses reported, dropout did not interact with the effects of the factors in question; therefore, participants who returned to the following sessions did not differ significantly from those who did not return in terms of the effect of ambiguity or ambiguity source. The one exception to this is reported with the appropriate analysis (see footnote 3).

Delayed Dutch-to-English production test: Second session

A second Dutch-to-English translation production test was administered at the beginning of the second session, before participants were trained again. It therefore reflects performance after a two-day delay. Thirty-seven participants returned for the second session, but because of technical problems, production data from four participants were lost. Therefore, data from 33 participants are included in the analyses that follow.

Accuracy

As predicted, ambiguous items were produced less accurately ($M = .48$) than unambiguous items ($M = .64$), $F_1(1,32) = 33.83$, $MSE = .01$, $p < .01$; $F_2(1,70) = 7.93$, $MSE = .06$, $p < .01$. As in the immediate test during the first session, increasing the number of repetitions of the ambiguous items during training was not sufficient to alleviate their disadvantage; ambiguous Dutch items repeated just as often as unambiguous items were still produced less accurately ($M = .53$) than unambiguous items ($M = .64$), $F_1(1,32) = 15.87$, $MSE = .01$, $p < .01$,

https://doi.org/10.1017/S1366728909990411 Published online by Cambridge University Press

but the effect was only marginally significant by items, $F_2(1,70) = 3.65$, $MSE = .04$, $p = .06$.

Furthermore, source of ambiguity had a reliable effect by participants, $F_1(2,64) = 7.59$, $MSE = .02$, $p < .01$, and a marginally-significant effect by items, $F_2(2,69) = 2.62$, $MSE = .04$, $p = .08$. Planned comparisons showed that form-ambiguous items were translated less accurately ($M = .50$) than unambiguous items ($M = .64$), $t(32) = 4.38$, $p < .01$, and marginally less accurately than meaning-ambiguous items ($M = .57$), $t(32) = 1.71$, $p = .10$. In addition, meaning-ambiguous items were translated less accurately than unambiguous items, $t(32) = 2.04$, $p < .05$.

Reaction times

Trimming procedures identical to those used for data from the first session led to the removal of 3.9% of the data. Although translation-ambiguous items were still produced more slowly ($M = 1732$) than unambiguous items ($M = 1630$), the effect of translation ambiguity was only marginally significant, $F_1(1,32) = 2.97$, $MSE = 57788$, $p = .09$; $F_2(1,69) = 2.20$, $MSE = 224865$, $p > .10$. Moreover, the effect of ambiguity was not significant when the number of repetitions of Dutch ambiguous items was equated to that of unambiguous items, $Fs < 1$. Source of ambiguity also did not reliably affect translation latency, $Fs < 1$.

To summarize the production data from the second session, the effect of translation ambiguity was evident in the accuracy data even after a two-day delay. Ambiguous items were still produced less accurately than unambiguous items, even when repeated just as often during training; by contrast, the reaction time decrement for ambiguous items did not hold when equated for number of presentations. In addition, form-ambiguous and meaning-ambiguous items were produced less accurately than unambiguous items, with a marginally-significant benefit for meaning-ambiguous items.

Longer-term retention Dutch-to-English production test: Third session

The third Dutch-to-English production test was administered at the beginning of the third session, and thus reflects retention after a delay of two and a half weeks. Note, however, that data from only 17 participants were available for the third-session analyses due to participant dropout. The results may therefore reflect reduced power.

Accuracy

Ambiguous items were produced less accurately ($M = .42$) than unambiguous items ($M = .52$) by participants, $F_1(1,16) = 12.78$, $MSE = .01$, $p < .01$; $F_2 < 1$. When ambiguous items were provided with more training to equate their familiarity with that of unambiguous items, the effect did not reach significance, $F_1(1,16) = 1.48$, $MSE = .01$, $p > .10$; $F_2 < 1$, but was in the same direction ($M = .48$ and .53, respectively). The effect of ambiguity source was not significant, $Fs < 1$.

Reaction times

Trimming procedures identical to those used on data from the first and second sessions led to the removal of 2.5% of the data. The effect of translation ambiguity was not significant, both when ambiguous Dutch items were trained half as often and when they were trained as often as unambiguous items, $Fs < 1$. Source of ambiguity also did not reliably affect reaction time, $Fs < 1$.

To summarize, the results of the Dutch-to-English translation production test from the third session suggest that the pattern observed during the first two sessions holds even after a two-and-a-half-week delay. The effect of translation ambiguity was significant in the accuracy data, however only in the analysis by participants and only when unambiguous items were repeated more often. With only 17 participants contributing data to the above analyses, it is possible that between-participants variability can explain why these effects did not reach significance in the item analyses.

First translation recognition test: Second session

The first recognition test was administered following a Dutch-to-English production test and one training cycle on the second session. Note that accuracy is expected to be higher on this task than on production tasks because chance performance is 50%. Data from 37 participants are included in the analyses that follow.

Accuracy

Ambiguous items were recognized less accurately ($M = .94$) than unambiguous items ($M = .97$) by participants and marginally so by items, $F_1(1,36) = 5.29$, $MSE = .004$, $p < .05$; $F_2(1,70) = 2.85$, $MSE = .01$, $p = .10$. When ambiguous items were trained just as often as unambiguous items, the ambiguous items were at a disadvantage according to the analysis by participants, $F_1(1,36) = 5.33$, $MSE = .003$, $p < .05$; $F_2(1,70) = 1.70$, $MSE = .01$, $p > .10$; ambiguous items ($M = .95$) were still recognized less accurately than unambiguous items ($M = .97$). The effect of ambiguity source was

https://doi.org/10.1017/S1366728909990411 Published online by Cambridge University Press

not significant, $F_1(2,72) = 1.39$, $MSE = .01$, $p > .10$; $F_2(2,69) = 1.23$, $MSE = .01$, $p > .10$.

Reaction times

Based on the distribution of the data, reaction times shorter than 600 ms or longer than 4000 ms were removed as outliers, along with data points more than 2.5 standard deviations above or below each participant's mean for correct responses. These procedures resulted in the exclusion of 4.9% of the data. Because words from different pairs were used as foils for the "no" trials, analyses were performed on correct "yes" trials only (see also Talamas, Kroll & Dufour, 1999).

Translation-ambiguous items were recognized significantly more slowly in the analysis by items ($M = 1457$ by items) than unambiguous items ($M = 1326$ by items), $F_1(1,36) = 2.69$, $MSE = 51747$, $p > .10$; $F_2(1,70) = 6.01$, $MSE = 46047$, $p < .05$. Again, repeating ambiguous items just as often during training did not alleviate their disadvantage; ambiguous items ($M = 1401$) were still recognized more slowly than unambiguous items ($M = 1234$), $F_1(1,36) = 16.03$, $MSE = 32427$, $p < .01$, although the effect was not reliable by items, $F_2(1,70) = 2.46$, $MSE = 57133$, $p > .10$.

Further, the effect of ambiguity source was significant by participants, $F_1(2,68) = 7.98$, $MSE = 144757$, $p < .01$, but not by items, $F_2(2,69) = 2.21$, $MSE = 56390$, $p > .10$ (see Figure 2). Planned comparisons revealed that form-ambiguous items were responded to significantly more slowly ($M = 1534$) than meaning-ambiguous items ($M = 1304$), $t(34) = 2.37$, $p < .05$, and significantly more slowly than unambiguous items ($M = 1242$), $t(36) = 3.93$, $p < .01$, but that meaning-ambiguous items did not differ significantly from unambiguous items, $t(34) = 1.16$, $p > .10$.

To summarize, although significant only in the analyses by items or by participants, translation-ambiguous items were recognized less accurately and more slowly than unambiguous items. Furthermore, the recognition data again suggest that increasing the familiarity with ambiguous items by repeating them just as often during training is not sufficient to alleviate their learning disadvantage. The detailed comparison of meaning-ambiguous and form-ambiguous items revealed that form-ambiguous items suffer a learning disadvantage as compared to unambiguous items in the latency data.

Second recognition test: Third session

The second translation recognition test was administered following a Dutch-to-English production test in the third session. Data from 17 participants were included in the following analyses, and the results may therefore reflect reduced statistical power.

Accuracy

The effect of translation ambiguity on accuracy was reliable by participants, $F_1(1,16) = 7.31$, $MSE = .01$, $p < .05$; $F_2(1,62) = 2.26$, $MSE = .06$, $p > .10$, with ambiguous items being recognized less accurately ($M = .84$) than unambiguous items ($M = .91$).[3] Furthermore, as in the first recognition test, repeating ambiguous items as often as unambiguous items during training did not alleviate their disadvantage. Ambiguous items were still recognized less accurately ($M = .85$) than unambiguous items by participants ($M = .91$), $F_1(1,16) = 7.00$, $MSE = .01$, $p < .05$; $F_2(1,70) = 2.76$, $MSE = .06$, $p > .10$.

The effect of ambiguity source was significant by items, $F_1(2,32) = 1.98$, $MSE = .04$, $p > .10$; $F_2(2,69) = 4.37$, $MSE = .05$, $p < .05$. Planned comparisons revealed that form-ambiguous items were recognized marginally less accurately ($M = .81$) than unambiguous items ($M = .91$), $t(16) = 1.93$, $p = .07$, but that meaning-ambiguous items ($M = .88$) did not differ significantly from either, ps $> .10$.

Reaction times

Identical trimming procedures to those employed for the first recognition test were used and led to the removal of 3.8% of the data. The effects of translation ambiguity, ambiguity source, and ambiguity when comparing ambiguous items repeated as often as unambiguous items were not reliable, Fs < 1.

To summarize, the accuracy data from the second recognition test demonstrate a disadvantage for ambiguous items, which does not go away with additional repetition of the ambiguous items. The accuracy data also highlighted a learning disadvantage for form-ambiguous items. However, data from only 17 participants were available for these third session recognition analyses, and indeed none of the reaction time effects reached significance, likely due to reduced power.

English-to-Dutch production test: Third session

The English-to-Dutch production test was administered at the end of the third session. Data from 17 participants are included in the analyses that follow.

[3] This effect significantly interacted with dropout, $F_1(1,35) = 5.32$, $MSE = .002$, $p < .05$. Examination of the means revealed, however, that ambiguous items that were repeated as often during training were recognized numerically less accurately than unambiguous items, for both participants who returned to the third session of testing ($M = .93$, $M = .98$) and for those who did not, ($M = .96$, $M = .97$).

https://doi.org/10.1017/S1366728909990411 Published online by Cambridge University Press

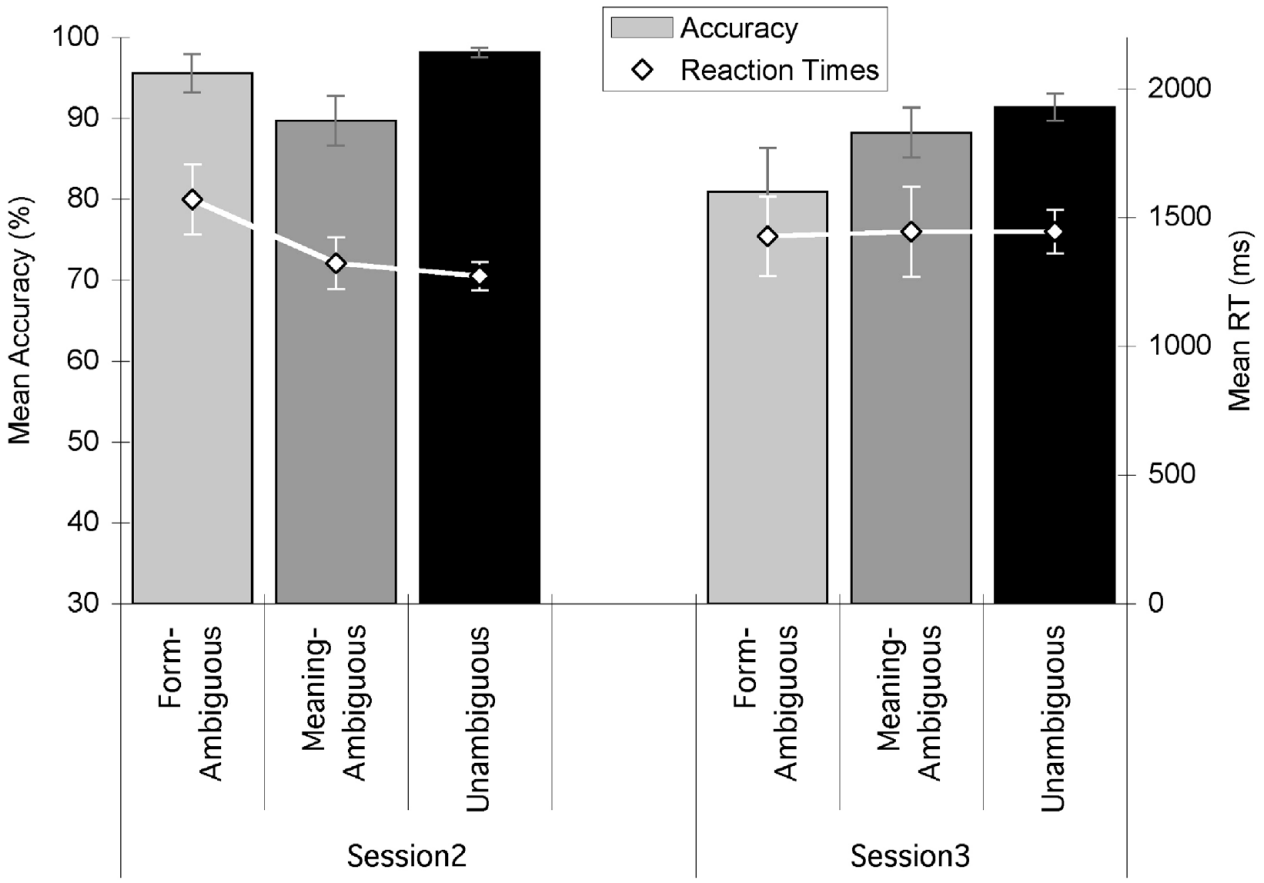

Figure 2. Accuracy and reaction time data from the translation recognition test in the second and third sessions. Data are shown as a function of the type of item (form-ambiguous, meaning-ambiguous, or unambiguous) in the equated repetition condition.

Accuracy

None of the effects were significant in the accuracy analyses, $ps > .10$.

Reaction times

Based on the distribution of the data, reaction times shorter than 300 ms or longer than 7000 ms, or that fell more than 2.5 standard deviations from each participant's mean reaction time for correct trials were removed from the analyses. These procedures resulted in the removal of 3.5% of the data. Analyses were performed on correct trials only.

The effect of translation ambiguity was not reliable by participants or by items, either when ambiguous items were presented less often than, $Fs < 1$, or as often as unambiguous items, $ps > 10$. However, the effect of ambiguity source was reliable by participants, $F_1(2,28) = 6.78$, $MSE = 109433$, $p < .01$; $F_2(2,64) = 1.29$, $MSE = 618459$, $p > .10$. Planned comparisons revealed that form-ambiguous items were translated significantly more slowly ($M = 2101$) than meaning-ambiguous ($M = 1656$), $t(14) = 4.47$, $p < .01$, and unambiguous ($M = 1895$) items, $t(14) = 2.45$, $p < .05$. The difference between meaning-ambiguous and unambiguous items was not significant, $t(16) = 1.35$, $p > .10$.

To summarize, the English-to-Dutch production test did not give rise to a robust ambiguity effect. Nonetheless, with only 17 participants the reaction time data still suggest that form-ambiguous items suffer a disadvantage relative to meaning-ambiguous and unambiguous items.

Comparing performance across tests and sessions

Several recognition and production tests were used to track performance over sessions. The pattern of results clearly demonstrates superior performance in recognition over production tasks, particularly with respect to accuracy, which is to be expected because chance-level performance would be 50% in the recognition tasks. Furthermore, translation from English to Dutch was performed less accurately than translation from Dutch to English,

https://doi.org/10.1017/S1366728909990411 Published online by Cambridge University Press

consistent with some previous studies (e.g., Kroll & Stewart, 1994).

To examine potential changes in performance across sessions, we analyzed the data for the 15 participants who contributed data across all sessions and in all conditions in the analysis by participants. These analyses are therefore limited in their statistical power. In the analysis by items, we were able to maximize statistical power by including data from all participants who participated in each session. We focused these analyses on accuracy, which gave rise to the most robust effects. The analysis of the data from the Dutch-to-English translation task demonstrates that participants' accuracy decreased across sessions, reflecting the fact that at later sessions, the tests measure retention because they did not follow training, $F_1(2,28) = 19.22$, $MSE = .04$, $p < .01$; $F_2(2,138) = 59.77$, $MSE = .02$, $p < .01$. There was no differential decay as a function of ambiguity or ambiguity source for the production test. In the recognition test, we again found that accuracy decreased from session two to session three, $F_1(1,16) = 17.79$, $MSE = .01$, $p < .01$; $F_2(1,69) = 14.32$, $MSE = .03$, $p < .01$. Moreover, there was a significant effect of ambiguity source in the analysis by items such that accuracy was highest for unambiguous items and lowest for form-ambiguous items, with meaning-ambiguous items in between, $F_1(2,32) = 1.99$, $MSE = .04$, $p > .10$; $F_2(2,69) = 4.40$, $MSE = .03$, $p < .05$. Importantly, there was differential decay across sessions such that accuracy for form-ambiguous items decreased more than for meaning-ambiguous items and for unambiguous items, $F_1(2,32) = 2.82$, $MSE = .02$, $p = .10$; $F_2(2,69) = 3.82$, $MSE = .03$, $p < .05$.

Discussion

In this study, we examined whether there is an ambiguity disadvantage in adult L2 word learning, and further whether the mapping of words-to-meanings PER SE influences this disadvantage. Across different tests and time points, we demonstrated that ambiguous words are harder to learn than unambiguous words. Specifically, ambiguous words were translated from Dutch to English more slowly and less accurately than unambiguous words immediately after learning, and this accuracy decrement remained even after a two-day delay. These ambiguous words were also recognized more slowly and less accurately, and showed some increase in reaction time when translated from English to Dutch. The ambiguity disadvantage was generally not alleviated by training the ambiguous items as often as the unambiguous items. Furthermore, the ambiguity disadvantage was more pronounced for the form-ambiguous items (two Dutch translations corresponding to the same meaning) than the meaning-ambiguous items (each translation corresponding to a different meaning of an ambiguous English word); the form-ambiguous items suffered the largest accuracy decay over sessions in the recognition task, further demonstrating the difficulty they pose for learning.

There are several possible reasons why one-to-many mappings such as those reflected by translation ambiguity pose difficulties. One possible mechanism is active competition between multiple translations. Interactive Activation (IA) models (e.g., McClelland & Rumelhart, 1981) assume a competitive inhibitory interaction between connected representations. In the case of words with more than one translation, one word is connected to two translations in the other language, and competition between these different alternatives is likely to create interference, whereas no such competition is expected for words with only one translation. Thus, the difficulty for adults when they try to learn ambiguous words may be rooted in this competition between alternatives. Likewise, Doherty (2004) suggested that the children in his study were poor at learning new meanings for pseudo-homonyms because they had difficulty suppressing the primary meaning of the homonyms, suggesting that competition from the primary meaning contributed to the observed effects.

A fan-type explanation could also provide a mechanism for the difficulty posed by multiple translations. In particular, the FAN EFFECT refers to the increase in latency to recall a concept as more facts are associated with that concept (e.g., Anderson, 1974). This phenomenon presumably reflects the decrease in associative strength from the concept to each fact, because as more facts are linked to a concept, the probability that a given fact will occur in the presence of the concept decreases (see Anderson & Reder, 1999, for discussion). Thus, in the case of translation ambiguity, the mapping of one word to two translations may decrease the associative strength of each translation pair.

Although both form- and meaning-ambiguous items could be subject to the same competition or fan-type effect, one-to-many mappings may be manifested differently for these two types of words. Indeed, form-ambiguous words suffered a greater disadvantage than meaning-ambiguous words in several tasks. Thus, learning two translations to the exact same meaning is especially difficult. In natural language learning, it is rarely the case that two word forms correspond to the exact same meaning in the manner implemented in the current study. It is more often the case that proficient language users are aware of slight meaning nuances or register differences between near-synonyms. However, these subtle meaning differences need to be acquired, and beginning L2 learners likely lack the experience to allow them to learn and appropriately

apply these meaning nuances. Moreover, many textbooks for beginning classroom instruction do not make these distinctions salient. Therefore, the exact-synonymy we created in the current study for form-ambiguous items, by presenting two Dutch translations with the exact same meaning definition, likely reflects the challenges faced by beginning L2 learners.

Meaning-ambiguous items were learned more easily than form-ambiguous items, despite the existence of translation ambiguity for both. This pattern of results would be predicted if one assumes that because ambiguous English words have one word form but two meanings, it is possible to maintain a one-to-one mapping, perhaps by directly mapping a Dutch translation to each meaning representation. The ability of beginning L2 learners to map L2 words directly to meaning representations is a central issue in models of the bilingual lexicon. The Revised Hierarchical Model (Kroll & Stewart, 1994), for example, suggests that beginning classroom learners typically access the meaning of L2 words via their L1 translations (but see e.g., Dufour & Kroll, 1995; Duyck & de Houwer, 2008, for a different pattern of results), and that the ability to map L2 word forms directly to meaning increases with proficiency (e.g., Kroll & Stewart, 1994; Talamas et al., 1999). The current study suggests that presenting learners with a meaning definition during training is a useful way to allow learners to more directly link L2 words to their meanings. Furthermore, directly mapping L2 words to their meanings appears to assist the learner in reducing the challenges of learning translation-ambiguous words.

Although meaning-ambiguous items had an advantage relative to form-ambiguous items on some measures, learning meaning-ambiguous items still appears to be more difficult than learning unambiguous items, although the numerical difference between them did not always reach significance. This lack of significance may be rooted in the variability of the meaning similarity of the meaning-ambiguous items. In particular, some ambiguous words have very unrelated meanings (e.g., the small coins and alteration meanings of *change*), and these two meanings were most likely already distinguished for our native English-speaking participants. Some ambiguous words, however, encompass two related meanings, or senses (e.g., *people*, denoting the body of citizens of a state or country, and more than one person), and these may not necessarily have a separate meaning representation for each sense (but see Klein & Murphy, 2001). Recent within-language ambiguity research has emphasized the importance of the degree of relatedness between the meanings or senses of an ambiguous word in explaining ambiguity effects (e.g., Armstrong & Plaut, 2008; Klepousniotou & Baum, 2007; Rodd et al., 2002). Thus, it may be important to consider the degree of relatedness of the two meanings (or definitions) of meaning-ambiguous words.

To examine the relatedness of the two meanings of each meaning-ambiguous item in our stimulus set, we collected normative data from a group of 27 monolingual English speakers who rated the semantic relatedness of the two definitions of each meaning-ambiguous word, on a scale from 1 (completely different) to 7 (exactly the same). These norms indicate that indeed our meaning-ambiguous items varied considerably in the degree of relatedness of their meanings ($M = 3.9$, range 1.5–5.8). Because our items were not chosen to examine the impact of meaning relatedness on learning, and were not matched as such, more detailed examination of this interesting issue is not possible with the current data. Nonetheless, the variability in the meaning relatedness of the items likely influenced the ease with which learners could map a translation uniquely to each meaning. Future studies will investigate this issue more directly.

The ambiguity in the current study was such that two Dutch translations corresponded to the same English word. There was therefore no output competition between two alternatives in the Dutch-to-English translation production task. Nonetheless, a robust ambiguity disadvantage was observed in this task. Thus, our findings extend previous research that showed that translation ambiguity impacts bilingual language processing when output competition is present (e.g., Tokowicz & Kroll, 2007), and show that translation ambiguity also influences learning.

Lastly, the current study shows that the ambiguity disadvantage goes beyond the decrement due to less-frequent exposure to translation-ambiguous items. With the exception of the Dutch-to-English production test administered during the last session (which had low statistical power), we found that repeating ambiguous items just as often during training was not sufficient to alleviate their accuracy disadvantage, and that one-to-many mapping, PER SE, creates challenges for learning.

In sum, the current study shows that translation ambiguity poses difficulty for adult L2 learning. Specifically, translation-ambiguous words, and particularly those that map onto a single meaning, were produced less accurately and more slowly than unambiguous pairs, both immediately after training and following a delay. These ambiguous pairs were further recognized less accurately and more slowly as correct translations. These learning decrements were not alleviated by repeating the ambiguous pairs more often during training. Because ambiguity is extremely prevalent in language, it can pose considerable difficulty for L2 learners. Therefore, it would be advantageous to identify teaching methods that are specifically tailored to these ambiguities, to provide learners with a better starting point when they try to communicate in L2.

Appendix. Complete set of stimuli and definitions by condition

Condition	English word	Definition(s)	Dutch translation(s)
Form-ambiguous	size	1. how large something is	1. grootte
			2. maat
Form-ambiguous	watch	1. a small portable timepiece typically worn on the wrist	1. horloge
			2. kijken
Form-ambiguous	sky	1. the region of the clouds or the upper air	1. lucht
			2. hemel
Form-ambiguous	decision	1. a position or opinion or judgment reached after consideration	1. beslissing
			2. besluit
Form-ambiguous	education	1. the activities of teaching or instructing that impart knowledge or skill	1. onderwijs
			2. opleiding
Form-ambiguous	nurse	1. a person educated and trained to care for the sick or disabled	1. verpleegster
			2. zuster
Form-ambiguous	case	1. an occurrence of something	1. geval
			2. zaak
Form-ambiguous	boot	1. protective footwear that covers the whole foot and lower leg	1. laars
			2. schoen
Form-ambiguous	sin	1. a transgression of a religious or moral law, especially when deliberate	1. zonde
			2. zondigen
Form-ambiguous	circle	1. a closed shape consisting of all points at a given distance from a center point within it	1. cirkel
			2. rondje
Form-ambiguous	attention	1. concentration of the mental powers upon an object	1. aandacht
			2. attentie
Form-ambiguous	influence	1. a power to affect persons or events especially power based on prestige, etc.	1. invloed
			2. influentie
Meaning-ambiguous	change	1. the result of alteration or modification	1. verandering
		2. coins of small denomination	2. wisselgeld
Meaning-ambiguous	interest	1. a fixed charge for borrowing money	1. rente
		2. a sense of concern with and curiosity about someone or something	2. interesse
Meaning-ambiguous	sheet	1. bed linen consisting of a large rectangular piece of cloth	1. laken
		2. an individual piece of paper used for writing or printing	2. blad
Meaning-ambiguous	iron	1. a silvery-white magnetic metallic element that rusts readily	1. ijzer
		2. to press and smooth with a heated appliance	2. strijken
Meaning-ambiguous	part	1. a portion of something	1. deel
		2. to force, take, or pull apart	2. verdelen
Meaning-ambiguous	people	1. the body of citizens of a state or country	1. volk
		2. more than one person	2. mensen
Meaning-ambiguous	triangle	1. a closed shape with three sides and three angles	1. driehoek
		2. a percussion instrument consisting of a metal bar bent in the shape of an open triangle	2. triangel
Meaning-ambiguous	wood	1. the hard substance under the bark of trees used to make things	1. hout
		2. the trees and other plants in a large densely wooded area	2. bos
Meaning-ambiguous	public	1. the community or the people as a whole	1. publiek
		2. open to all people	2. openbaar
Meaning-ambiguous	smell	1. the ability to perceive scents through the nose	1. reuk
		2. to emit an unpleasant odor	2. ruiken

https://doi.org/10.1017/S1366728909990411 Published online by Cambridge University Press

Appendix. Continued

Condition	English word	Definition(s)	Dutch translation(s)
Meaning-ambiguous	block	1. a three-dimensional shape with six square or rectangular sides 2. to prevent access or progress	1. blok 2. blokkade
Meaning-ambiguous	proposal	1. something offered or suggested, such as a plan 2. an offer of marriage	1. voorstel 2. aanzoek
Unambiguous	figure	form or shape, as determined by outlines or exterior surfaces	figuur
Unambiguous	height	distance from the base of something to its top	hoogte
Unambiguous	abuse	cruel or inhumane treatment	misbruik
Unambiguous	solution	a statement that solves a problem or explains how to solve the problem	oplossing
Unambiguous	science	systematic knowledge of the physical or material world gained through observation and experimentation	wetenschap
Unambiguous	captain	an officer with a rank below major but above lieutenant	kapitein
Unambiguous	field	an expanse of open or cleared ground	veld
Unambiguous	arrow	a mark with a pointed end used to indicate a direction or relation	pijl
Unambiguous	discovery	a breakthrough or finding that is uncovered	ontdekking
Unambiguous	bird	warm-blooded egg-laying vertebrate characterized by feathers and wings	vogel
Unambiguous	hospital	a health facility where patients receive treatment	ziekenhuis
Unambiguous	result	the consequence of a particular action, operation, or course; an outcome	resultaat
Unambiguous	night	the period of darkness between sunset and sunrise	nacht
Unambiguous	question	a request for information that calls for a reply	vraag
Unambiguous	beauty	the qualities that give pleasure to the senses	schoonheid
Unambiguous	paint	a dye or pigment used as a coating to protect or decorate a surface	verf
Unambiguous	army	a permanent organization of military land forces	leger
Unambiguous	boy	a youthful male person	jongen
Unambiguous	age	how old something is	leeftijd
Unambiguous	name	a word used to refer to a person or thing	naam
Unambiguous	time	the system of those sequential relations that any event has to any other, as past, present, or future	tijd
Unambiguous	bible	a book of sacred writings	bijbel
Unambiguous	butterfly	diurnal insect typically having a slender body with knobbed antennae and broad colorful wings	vlinder
Unambiguous	window	a framework of wood or metal that contains a glass windowpane and is built into a wall or roof to admit light or air	raam

References

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, 6, 451–474.

Anderson, J. R., & Reder, L. M. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General*, 128, 186–197.

Armstrong, B. C., & Plaut, D. C. (2008). Settling dynamics in distributed networks explain task differences in semantic ambiguity effects: Computational and behavioral evidence. Presented at the 30th Annual Conference of the Cognitive Science Society, Hillsdale, NJ.

Balota, D. A., Cortese, M. J., Hutchison, K. A., Neely, J. H., Nelson, D., Simpson, G. B., & Treiman, R. (2002). The English Lexicon Project: A web-based repository of descriptive and behavioral measures for 40,481 English words and nonwords. Washington University, http://elexicon.wustl.edu/ (retrieved August 14, 2007).

Boada, R., Sánchez-Casas, R., García-Albea, J. E., Gaviln, J. M., & Ferr, P. (2009). Effect of number of translations and cognate status in translation recognition performance of proficient bilinguals. Presented at the Seventh International Symposium on Bilingualism, Utrecht, The Netherlands.

Bogaards, P. (2001). Lexical units and the learning of foreign language vocabulary. *Studies in Second Language Acquisition*, 23, 321–343.

Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: Infants' language experience influences the development of a word-learning heuristic. *Developmental Science*, 12, 815–823.

Degani, T., Prior, A., & Tokowicz, N. (in press). Bidirectional transfer: The effect of sharing a translation. *European Journal of Cognitive Psychology*.

https://doi.org/10.1017/S1366728909990411 Published online by Cambridge University Press

Degani, T., & Tokowicz, N. (in press). Semantic ambiguity within and across languages: An integrative review. *Quarterly Journal of Experimental Psychology*.

Doherty, M. J. (2004). Children's difficulty in learning homonyms. *Journal of Child Language*, 31, 203–214.

Dufour, R., & Kroll, J. F. (1995). Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation. *Memory and Cognition*, 23, 166–180.

Duyck, W., & de Houwer, J. (2008). Semantic access in second-language visual word processing: Evidence from the semantic Simon paradigm. *Psychonomic Bulletin and Review*, 15, 961–966.

Electronic commerce. (2007). *Dictionary.com.* http://dictionary.com (retrieved October 22, 2007).

Elston-Güttler, K. E., & Friederici, A. D. (2005). Native and L2 processing of homonyms in sentential context. *Journal of Memory and Language*, 52, 256–283.

Frenck-Mestre, C., & Prince, P. (1997). Second language autonomy. *Journal of Memory and Language*, 37, 481–501.

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology*, 28, 99–108.

Jiang, N. (2002). Form–meaning mapping in vocabulary acquisition in a second language. *Studies in Second Language Acquisition*, 24, 617–637.

Jiang, N. (2004). Semantic transfer and its implications for vocabulary teaching in a second language. *The Modern Language Journal*, 88, 416–432.

Klein, D. E., & Murphy, G. L. (2001). The representation of polysemous words. *Journal of Memory and Language*, 45, 259–282.

Klepousniotou, E., & Baum, S. R. (2007). Disambiguating the ambiguity advantage effect in word recognition: An advantage for polysemous but not homonymous words. *Journal of Neurolinguistics*, 20, 1–24.

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

Lotto, L., & de Groot, A. M. B. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning*, 48, 31–69.

MacWhinney, B. (1997). Second language acquisition and the Competition Model. In A. M. B. de Groot & J. F. Kroll (eds.), *Tutorials in bilingualism: Psycholinguistic perspectives*, pp. 113–142. Mahwah, NJ: Lawrence Erlbaum.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, 20, 121–157.

Mazzocco, M. M. M. (1997). Children's interpretations of homonyms: A developmental study. *Journal of Child Language*, 24, 441–467.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception. Part 1: An account of basic findings. *Psychological Review*, 88, 375–407.

Prior, A., MacWhinney, B., & Kroll, J. F. (2007). Translation norms for English and Spanish: The role of lexical variables, word class, and L2 proficiency in negotiating translation ambiguity. *Behavior Research Methods*, 39, 1029–1038.

Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, 46, 245–266.

Sánchez-Casas, R., Buratti, B. S., & Igoa, J. M. (1992). Are bilingual lexical representations interconnected? Presented at the Fifth Conference of the European Society for Cognitive Psychology, Paris, France.

Storkel, H. L., & Maekawa, J. (2005). A comparison of homonym and novel word learning: The role of phonotactic probability and word frequency. *Journal of Child Language*, 32, 827–853.

Talamas, A., Kroll, J. F., & Dufour, R. (1999). From form to meaning: Stages in the acquisition of second-language vocabulary. *Bilingualism: Language and Cognition*, 2, 45–58.

Tokowicz, N. (2005). The role of proficiency in the influence of L1/L2 differences on L2 processing. Presented at symposium Cross-language Interaction at Different Levels of Proficiency: Psycholinguistic Studies of Late L2 Learners, chaired by J. G. van Hell & T. Dijkstra, held at the Fifth International Symposium on Bilingualism, Barcelona, Spain.

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

Tokowicz, N., Kroll, J. F., de Groot, A. M. B., & van Hell, J. G. (2002). Number-of-translation norms for Dutch–English translation pairs: A new tool for examining language production. *Behavior Research Methods, Instruments & Computers*, 34, 435–451.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in Second Language Acquisition*, 27, 173–204.

Tokowicz, N., Michael, E. B., & Kroll, J. F. (2004). The roles of study-abroad experience and working-memory capacity in the types of errors made during translation. *Bilingualism: Language and Cognition*, 7, 255–272.

Tokowicz, N., Michael, E. B., & Smith, C. J. (2007). Choices, choices, choices: The consequences of activating too many translation alternatives. Presented at the Forty-Eighth Annual Meeting of the Psychonomic Society, Long Beach, CA.

Tokowicz, N., Prior, A., & Kroll, J. F. (2009). Bilingual speech production depends on translation ambiguity. Ms., University of Pittsburgh.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28, 127–154.

Wilson, M. D. (1988). The MRC psycholinguistic database: Machine readable dictionary, version 2. *Behavior Research Methods, Instruments & Computers*, 20, 6–11.

WordNet 3.0, Princeton University, 2006.