Together or apart: Learning of translation-ambiguous words*

TAMAR DEGANI
University of Haifa
ALISON M. TSENG
University of Pittsburgh
NATASHA TOKOWICZ
University of Pittsburgh

In a multiple-session training study, native English speakers learned foreign Dutch vocabulary items that mapped to English either in a one-to-one way (translation-unambiguous) or in a one-to-many way (translation-ambiguous), such that two Dutch words corresponded to a single English translation. Critically, these two translation-ambiguous Dutch words were taught on consecutive trials in the same session, or were presented separately, such that each word was taught in a separate session. Translation-ambiguous words were produced and recognized substantially less accurately than translation-unambiguous words on tests administered one and three weeks after training. An ambiguity advantage emerged, however, in a free-recall test. Interestingly, teaching both translations together led to superior performance over teaching them in separate sessions, in which case the translation learned first enjoyed a considerable advantage over that learned second. These findings underscore the importance of order of acquisition in second-language vocabulary learning, and have practical implications for language instruction.

Keywords: foreign vocabulary learning, translation ambiguity, order of acquisition

Introduction

Vocabulary is an important component of language learning, because it supports efficient spoken and written communication. Indeed, word knowledge has been linked to reading comprehension in both the first language (L1) and the second language (L2) (e.g., Pasquarella, Gottardo & Grant, 2012; Perfetti, 2007). Further, because adults are generally better at learning L2 vocabulary than grammatical and morphosyntactic structures, they may wish to capitalize on this ability to support increased proficiency in L2 (see e.g., Morgan-Short, Sanz, Steinhauer & Ullman, 2010, for ERP evidence of semantic processing of morphosyntactic violations among L2 learners). Notably, however, even vocabulary learning does not proceed without challenges for adults, with profound and lasting difficulty for particular types of words. For instance, abstract words are harder to learn than concrete words (for review, see De Groot & Van Hell, 2005), and these words are processed differently even by

proficient bilinguals (e.g., De Groot, Dannenburg & Van Hell, 1994; but see Tokowicz & Kroll, 2007).

Of particular relevance to the current study, TRANSLATION-AMBIGUOUS WORDS, which have more than one translation across languages, create enduring challenges for learning (Degani & Tokowicz, 2010). Degani and Tokowicz taught native English speakers Dutch translations for English words. In the translation-unambiguous condition, a single Dutch word corresponded to a single English word. In the translation-ambiguous condition, two Dutch words corresponded to a single English word. Speakers had more difficulty learning translation-ambiguous than translation-unambiguous words. This difficulty was reflected in translation production and translation recognition tasks, and was evident even two weeks after initial learning. This TRANSLATION-AMBIGUITY DISADVANTAGE was not simply due to lesser frequency for each of the two Dutch translations in the translation-ambiguous condition, because it held even when frequency was matched. Further, the translation-ambiguity disadvantage was larger for two Dutch synonyms that map to a single English word (e.g., *lucht* and *hemel* for the English word *sky*; referred to as FORM AMBIGUITY) than for two Dutch translations that each map to a different meaning of an ambiguous English word (e.g., *verandering* and *wisselgeld* to denote the 'alteration' and 'small coins' meanings of the English word *change*, respectively; referred to as

Address for correspondence:
Tamar Degani, The Institute of Information Processing and Decision Making, University of Haifa, 199 Aba Khoushy Ave, Mount Carmel, Haifa, Israel 3498838
Tdegani@research.haifa.ac.il

MEANING AMBIGUITY). Degani and Tokowicz suggested that mapping two lexical forms to a single undifferentiated meaning is the primary source of the observed learning difficulty, and, further, that this difficulty can be reduced when learners are able to maintain a one-to-one mapping by linking each translation to its appropriate meaning directly, as with meaning translation-ambiguous words.

Critically, translation-ambiguity is not a rare phenomenon. For instance, of a list of Dutch/English translations that had been used in previous studies, 25% had more than one translation from Dutch to English or the reverse (Tokowicz, Kroll, De Groot & Van Hell, 2002). Importantly, these items initially had been chosen under the assumption that they entailed direct one-to-one mappings across languages. Although it is difficult to estimate ambiguity in general from lists for psycholinguistic studies, the amount of ambiguity found in these lists (up to 55%; Prior, MacWhinney & Kroll, 2007) suggests that the difficulty associated with learning translation-ambiguous words may affect substantial portions of a learner's vocabulary. Moreover, translation-ambiguous words continue to pose difficulty with increased proficiency (for a review, see Tokowicz & Degani, 2010). Alleviating the difficulty at the early stages of learning may have long-lasting benefits as learners gain proficiency in L2.

In the current study, we test whether the observed difficulty in learning translation-ambiguous words can be alleviated by an instructional manipulation. In particular, in one condition, the two translations of a translation-ambiguous word were presented on consecutive trials in the same training session (TOGETHER CONDITION). For example, the learner would be told that *lucht* means *sky* on one trial, and that *hemel* means *sky* on the next trial. In the other condition, each translation was presented in a different training session (SEPARATE CONDITION). For example, the learner would be told that *lucht* means *sky* in the first training session and that *hemel* means *sky* on the second training session two days later. Further, when the two translations were presented in separate training sessions, we examined whether there is a difference between the translation learned first and the translation learned second (e.g., primacy effects in learning multiple labels for objects; Poepsel, Gerfen & Weiss, 2012).

Teaching both translations on consecutive trials in the same training session may be advantageous for several reasons. First, presenting both translations from the start may aid participants in establishing the appropriate one-to-many association structure, thus avoiding the need to change their initial mapping later in learning. Indeed, there is evidence to suggest that on the first encounter with a word, learners tend to create a single conjecture and seek its confirmation, thus requiring effort and additional evidence to change it if it is later found to be erroneous (Medina, Snedeker, Trueswell

& Gleitman, 2011). Second, the joint presentation of both translations may aid learners in differentiating them, drawing attention to discriminative attributes (Underwood, 1969), hence reducing potential competition between these two alternative translations (see similar suggestions regarding the joint recall of items in paired-associate learning paradigms; Postman & Gray, 1977).

In contrast, in typical classroom courses, often a single translation is learned first, and a second translation is withheld until later weeks or semesters. This is likely done under the assumption that it is better to allow the representation of the first translation to become fully stable before introducing additional complexities. As noted above, however, there is ample reason to suspect that introducing the relevant complexities early on may in fact lead to superior learning.

In one study, McDaniel, Neufeld and Damico-Nettleton (2001) examined learning of one-to-many mappings in a naturalistic context. Specifically, participants learned to use the icons of an alternative and augmentative communication (AAC) device. The device used a typical coding scheme in which each icon is associated with four words: a noun describing the referent of the icon, a category to which the referent belongs, and a verb and an adjective associated with that referent. The authors manipulated whether, during training, the icon was presented first followed by an oral production of one of its associates, or whether a word was first produced followed by the presentation of (pointing to) the icon. The former led to better retention but slower acquisition. Of relevance to the current investigation, in Experiment 1 learners were presented with pairings blocked by icon, such that all four words associated with a single icon were learned together. In contrast, in Experiment 2, training was spaced for each icon, such that the four words associated with each icon were not grouped together, and instead were presented according to their grammatical category (i.e., all nouns, all verbs, etc.). A comparison of the findings across the two experiments shows that blocking training according to icon produced better retention than interleaved training. These findings suggest that teaching multiple alternatives together may yield superior learning over teaching the alternatives separately. The current study was designed to examine this issue more directly, in relation to translation ambiguity.

In the present study, we contrasted learning of translation-unambiguous words with that of translation-ambiguous words, stemming from either meaning or form ambiguity (e.g., Degani & Tokowicz, 2010). The Revised Hierarchical Model of Translation Ambiguity (RHM-TA; see Figure 1; adapted from Eddington & Tokowicz, 2013; see also Kroll & Tokowicz, 2001) describes the way that the two types of translation-ambiguous words may be represented. This model is a modification of the Revised Hierarchical Model (Kroll & Stewart, 1994),
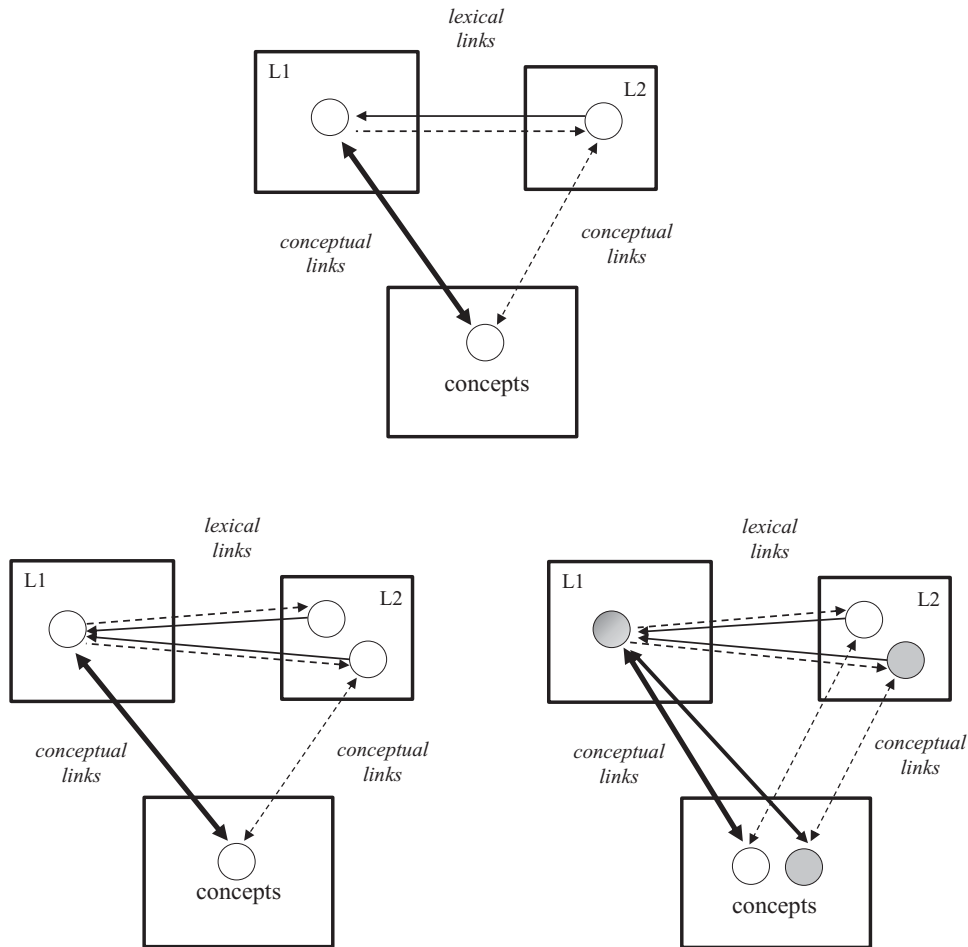
Figure 1. The Revised Hierarchical Model of Translation Ambiguity, adapted from Eddington and Tokowicz (2013). Translation-unambiguous (top), form-ambiguous (bottom left), meaning-ambiguous (bottom right).

which posits separate lexicons for the two languages and a shared conceptual store. The RHM-TA emphasizes that both types of translation-ambiguous words differ from unambiguous items in the mapping of translations at the lexical level. Whereas for translation-unambiguous items, a single L1 lexical form corresponds to a single L2 lexical form, in the case of translation-ambiguous items, a single L1 lexical form corresponds to two separate L2 lexical forms. For the form-translation ambiguous items, a one-to-many mapping also exists from a concept to two L2 forms. In contrast, the links from concept to L2 form are unambiguous in the case of meaning-ambiguous items, because each L2 lexical form has its own meaning. That is, native speakers already possess some differentiation of meaning for homographs such as 'change'. In our previous work, we have shown that when ambiguity exists at both the lexical and conceptual levels (i.e., form-ambiguous items), learning is especially difficult (Degani & Tokowicz, 2010). In the current study, we test whether different instructional methods alter the way in which these two types of ambiguity are learned. Although the

RHM-TA can explain differences in the processing of form and meaning translation-ambiguous words, it is agnostic with respect to the particular mechanisms that may make the learning of these words more or less difficult under different instructional methods.

If teaching both translations together to beginning learners helps to create the appropriate translational (lexical) links, with little regard to semantic mapping, we would expect this instructional method to operate similarly for form- and meaning-ambiguous items. Such a finding would follow suggestions that very beginning L2 learners rely less on conceptual links (Kroll & Stewart, 1994), or would suggest that for beginning learners our instructional method manipulation operates more at the lexical (rather than semantic) level. Alternatively, if teaching both translations together serves to create the appropriate links from concept to L2 lexical forms, we may expect stronger instructional method effects for form-ambiguous items, because for those words the links from concept to L2 lexical form are indeed ambiguous. For meaning-ambiguous items, in contrast, these links from

concept to L2 lexical forms are unambiguous and thus would benefit less from mechanisms that would pull the two semantic representations apart.

One-to-many mappings, at either the lexical or semantic level, may pose difficulty for learning due to active competition between the translations or because of reduced association strength due to a fan-type effect (for discussion see Degani & Tokowicz, 2010). In particular, because the two alternative L2 translations are linked to a single L1 lexical representation, inhibitory connections may develop between them, leading to active competition and interference during learning (e.g., McClelland & Rumelhart, 1981). Teaching both translations together would reduce competition because it draws attention to discriminative features (Underwood, 1969). Conceived of differently, the ultimate link structure in the case of translation ambiguity is one in which connection weights from the L1 word need to be divided between two links, one for each translation. Such reduced associative strength (weight) has been proposed to explain the fan-type effect in memory research (Anderson, 1974; Anderson & Reder, 1999) and in translation ambiguity (Degani & Tokowicz, 2010). If the learner starts out assigning the inappropriate weight for one translational link, wrongly assuming it is the only correct translation, he will later need to reassign the weights, and this process may take more effort and time (see also Medina et al., 2011).

We also examined whether the presentation order of the two translation-ambiguous Dutch words has consequences for later retention. The results of a recent word-training study suggest that the order of acquisition of foreign words influences their learning trajectory above and beyond their cumulative frequencies. In particular, Izura, Pérez, Agallou, Wright, Marín, Stadthagen-González and Ellis (2011) demonstrated that words taught early in learning yield faster responses in object naming, visual lexical decision, and semantic categorization tasks than words taught later in learning. Notably, in that study, items were unrelated. Here, we ask whether the first Dutch translation learned for a specific English word will be retained better than a second Dutch translation learned for the same English word. This type of design resembles classical paired-associate learning studies, in which participants are to learn first-list items A–B, and second-list items A–C, yielding the potential for proactive and retroactive interference.

In typical paired-associate learning tasks, participants are required to substitute one list of associations for the other, such that only the first-list items or the second-list items are considered correct on the test (for a review, see Anderson & Neely, 1996; Postman & Underwood, 1973). When first-list associations are recalled, A–C pairings are considered errors, and the results yield evidence for retroactive interference in comparison to a control condition (A–B, C–D) (e.g., Potts & Shanks, 2012; Torres,

Flashman, O'Leary & Andreasen, 2001). When second-list associations are recalled, A–B pairings are considered errors, and the results highlight proactive interference (e.g., Wahlheim & Jacoby, 2011). Participants are rarely asked to retain both sets of associations, as is done in the current study (see Postman & Parker, 1970).

Directly comparing recall of first-list items (A–B) and second-list items (A–C), Tulving and Watkins (1974) identified what they termed *The Priority Effect*, which occurs when recall of A–B items is higher than that of A–C items. They suggested that the priority effect stems from the lack of testing of the A–B list immediately prior to A–C learning (but see Potts & Shanks, 2012). In the current study, the second translation is presented in a separate training session in which the corresponding A–B association is not presented or tested. We therefore predict an advantage in learning for the translation learned first.

Importantly, however, the current study differs from traditional paired-associate learning in several ways. First, items were not taught to criterion. Second, as discussed above, learning in Session 2 was not intended to replace learning in Session 1, or vice versa, because learners were to maintain both sets of pairings for future recall and testing. Third, because we taught novice learners L2 vocabulary, they were to learn not only the association between A and B or A and C, but also had to create representations for the new items (B and C). Fourth, Tulving and Watkins' (1974) results suggest that it is not necessary to test particular A–B items to reduce the priority effect over A–C items, but rather that testing A–B items in general in conjunction with A–C learning is sufficient to reduce the effect (i.e., it is a list-wise rather than an item-specific effect). In the current study, participants were re-presented with some of the items from the first session during the second training session. We will examine whether an order of acquisition/priority effect is still manifested in the separate condition.

To summarize, using a within-participants design, the current study compared learning of translation-ambiguous and translation-unambiguous words. We expect an overall translation-ambiguity disadvantage (e.g., Degani & Tokowicz, 2010). Importantly, we further contrasted two instructional methods for teaching translation-ambiguous words. In the together condition, the two translations were presented on consecutive trials in both sessions. In the separate condition, each translation was taught only in either the first or second session. To rule out the influence of frequency of exposure, cumulative frequency was equated by teaching separate-condition translations twice as often as the together-condition translations in that training session (see Izura et al., 2011). We expect translation-ambiguous words to be learned better in the together condition than in the separate condition. To the extent that the instructional method manipulation involves differentiation of meaning, we would expect this

Table 1. *Language history questionnaire data for the final set of 28 participants.*

Measure	Average	SD
Age (years)	18.43	.96
Age began L2 (years)	14.28	1.02
Time studied L2 (years)	3.87	1.17
L1 reading ability	9.68	.48
L2 reading ability	3.78	1.50
L1 writing ability	9.50	.64
L2 writing ability	3.26	1.83
L1 conversation ability	9.93	.26
L2 conversation ability	3.33	2.04
L1 speech comprehension ability	9.86	.36
L2 speech comprehension ability	3.67	2.18

Notes: 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. L2 varied across participants but was never Dutch or German (6 participants listed French, 7 participants listed Spanish, 15 participants listed no L2).

effect to be stronger for form- vs. meaning-ambiguous items. In contrast, if it operates mainly at the lexical level, we would expect all translation-ambiguous words (form- and meaning-ambiguous alike) to be learned better when presented together vs. separately. Moreover, in the separate training sequence, we expect to observe gains for the translation learned first over the translation learned second.

Although the translation-ambiguity disadvantage could be found in either accuracy or reaction times, we expect larger effects in accuracy measures. This is based on the results of our previous translation-ambiguity training study (Degani & Tokowicz, 2010), and findings that non-native speakers demonstrate effects in accuracy for which native speakers demonstrate effects in reaction time (for discussion, see e.g., DeKeyser & Larson-Hall, 2005).

Method

Participants

Twenty-eight monolingual English speakers (eight males; mean age 18.5 years) with no prior exposure to Dutch or German participated in this experiment. They were recruited from a large north-eastern university in the United States, and received class credit for participating in Session 1, and either class credit or payment for participating in the remaining sessions. At the end of Session 1, participants completed a language history questionnaire (from Tokowicz, Michael & Kroll, 2004), designed to investigate L2 learning experiences (see Table 1). Data from 40 additional participants were

excluded because they did not complete all four sessions (17 participants), did not follow instructions (1 participant), or did not meet our recruitment criteria: were exposed to German (15 participants) or to languages other than English before age 12 (7 participants). All participants were right handed.

Design

A 2 ambiguity status (translation-unambiguous vs. translation-ambiguous) within-participants design was used. For translation-ambiguous items, a 2 ambiguity source (form vs. meaning) by 2 instructional method (together vs. separate) within-participants design was used. For items taught in the separate condition, a 2 ambiguity source (form vs. meaning) by 2 presentation order (first vs. second) within-participants design was used.

Stimuli

Stimuli were identical to those used by Degani and Tokowicz (2010), and consisted of 48 English words, presented with their Dutch translation(s) and a definition of their meaning(s) in English (see Table 2 for examples, and the appendix for the full list of items). Half of the English words had one translation into Dutch (translation unambiguous), and half had two translations into Dutch, either due to multiple meanings of the English word (meaning ambiguous) or due to near-synonymy in Dutch (form ambiguous). Because half of the English words had two translations in Dutch, the stimuli included a total of 72 Dutch words. An English definition was created for each of the Dutch words (see Degani & Tokowicz, 2010), resulting in two different definitions for the meaning-ambiguous English items, and one definition for the translation-unambiguous and form-ambiguous English items.

Stimulus characteristics are presented in Table 3. Translation-unambiguous items were matched to meaning and form-ambiguous items on English length, English log HAL frequency, and mean English naming reaction time and accuracy (all Fs < 1, except mean naming RT, p > .26, from Elexicon, Balota, Yap, Cortese, Hutchinson, Kessler, Loftis, Neely, Nelson, Simpson & Treiman, 2007), and on familiarity, concreteness, imageability, and age of acquisition of the English words (Fs < 1, from MRC database, Wilson, 1988). Further, translation-unambiguous items did not differ significantly from meaning or form-ambiguous items in Dutch length or in the form-similarity of each English-Dutch pair (a cognate-like measure, Tokowicz et al., 2002) (Fs < 1).

Twelve training versions were counterbalanced across participants.[1] Each version included 16 translation-ambiguous English words (eight form- and eight

[1] Due to the extended protocol, some participants dropped out and some data points are missing for the production tests due to technical problems. To maximize power, we included the maximum number

Table 2. *Example stimuli and definitions by condition.*

Condition	English word	Definition	Dutch translation(s)
Unambiguous	boy	1. a youthful male person	1. jongen
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
Meaning-ambiguous	sheet	1. bed linen consisting of a large rectangular piece of cloth 2. an individual piece of paper used for writing or printing	1. laken 2. blad

Table 3. *Stimulus characteristics.*

	Condition		
	Unambiguous	Form-ambiguous	Meaning-ambiguous
English word with Dutch translation(s)	arrow pijl	boot laars/schoen	change verandering/wisselgeld
English length (number of letters)	5.75 (1.75)	5.75 (2.38)	5.75 (1.55)
English log HAL frequency	6.54 (2.25)	6.88 (1.64)	6.42 (1.87)
Familiarity rating	534.04 (123.90)	517.67 (165.66)	480.33 (226.50)
Concreteness rating	435.00 (170.31)	399.92 (177.33)	405.75 (216.59)
Imageability rating	471.92 (170.55)	446.17 (180.68)	415.33 (211.74)
Age of acquisition rating	178.71 (180.92)	234.25 (195.41)	204.25 (159.47)
Mean naming reaction time	599.11 (40.11)	622.17 (38.91)	603.93 (40.13)
Mean naming accuracy	1.00 (.01)	1.00 (.01)	.99 (.02)
Average Dutch length (number of letters)	6.54 (2.25)	7.17 (2.25)	5.75 (2.26)
Average form similarity rating	2.79 (1.97)	2.32 (1.47)	2.22 (1.31)

Notes: Values are taken from Degani and Tokowicz (2010). Standard deviations are shown in parentheses. 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 Dutch–English bilinguals and were reported in Tokowicz et al. (2002).

meaning-ambiguous) presented with each of their Dutch translations and English definitions. For half of these, participants were trained in the together condition with the two Dutch translations on consecutive trials in Sessions 1 and 2 (order of translations was reversed across sessions for each participant),[2] and for the other half, participants were trained in the separate condition with one Dutch translation in Session 1 and the other Dutch translation in Session 2; order of translations was counterbalanced across participants. In addition, 32 English words were presented with only one Dutch translation and definition, to reduce the likelihood

that participants would develop strategies tailored to translation-ambiguity (e.g., Tokowicz & Kroll, 2007). Of these, 24 were real translation-unambiguous items, and eight were translation-ambiguous items presented with one translation only (across all training and testing sessions) that served as fillers in the analyses.[3]

Participants were trained in Sessions 1 and 2. During Session 1, each participant was presented with eight translation-ambiguous items with both Dutch translations, eight translation-ambiguous items with one Dutch translation (the second Dutch translation was presented during Session 2), and 32 translation-unambiguous items. Thus, a given participant was asked to learn 56 Dutch words in Session 1, and an additional 8 Dutch words in

---

of participants who completed all sessions; by doing so we were unable to maintain an equal number of participants in each of the 12 experimental versions. Analyses by items were therefore performed to verify that the effects were not driven by outlier items or versions.

[2] Because the order of translations was reversed for each participant between Session 1 and Session 2, the presentation order effect was only examined in the Separate condition, were it was instantiated consistently.

[3] These were initially designed to serve as within-item controls for the translation-ambiguous items. However, they were inadvertently presented with only one of their translations, and were therefore treated as fillers.

Table 4. *Training sequence by condition.*

Condition	# of English words	Dutch translations	Number of repetitions		
			Session 1	Session 2	Total
Unambiguous	24	a (only)	3	3	6
Together	8 (4 form, 4 meaning)	a	3	3	6
		b	3	3	6
Separate	8 (4 form, 4 meaning)	a	6	–	6
		b	–	6	6
Filler single	8 (4 form, 4 meaning)	a (only)	3	3	6
Total	48	64 (only 56 in each session)			

Table 5. *General procedure of training and testing.*

Session 1	Session 2 (2 days later)	Session 3 (1 week later)	Session 4 (2 weeks later)
Training cycle	Training cycle	**Free recall**	**Free recall**
Training cycle	Training cycle	**Dutch-to-English production**	**Dutch-to-English production**
Training cycle	Training cycle	Stroop	
Dutch-to-English production	Dutch-to-English production	**Translation recognition**	
Operation span		Simon	
Language history questionnaire		English-to-Dutch production	

Note. Bold measures reflect analyses that are reported in the text.

Session 2. In total, 64 Dutch words were learned in each version. This training sequence is presented in Table 4.

Procedure

The experiment consisted of four sessions, separated by two days, one week, and two weeks, respectively. Training took place in the first two sessions, and testing took place at the end of each training session to increase learning (e.g., Karpicke & Roediger, 2008) and during Sessions 3 and 4. The general procedure is summarized in Table 5. During training, participants were instructed to learn the Dutch words and their translations, and were told that some English words may have more than one Dutch translation. On each training trial, a fixation cross appeared in the center of the computer 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 a Dutch word with its English translation and definition for 8 seconds (following Lotto & De Groot, 1998). During each training session, participants completed three cycles of 56 training trials, each presented in a random order, with the exception that the two translations in the together condition were presented on consecutive trials. To equate the number of presentations of each Dutch translation, words in the separate condition were repeated twice in

each cycle because they were taught only in either Session 1 or Session 2. Therefore, in total, each translation was presented 6 times during the experiment.

Four different tests were administered, aimed at tapping different aspects of learning. The free recall test provides an unconstrained opportunity to retrieve the translation pairs in the absence of overtly presented cues. Participants are therefore able to choose any of several internally generated cues including a meaning, an L1 translation or an alternative L2 translation. It was administered first, before any additional information was provided by the tests themselves. To test participants' memory of the novel Dutch words more directly we used a Dutch-to-English translation production task (L2- to-L1 backward translation), often used with beginning learners (e.g., Kaushanskaya & Marian, 2009). Similarly, we tested recognition memory using a version of the translation-recognition task. Finally, we included an English-to-Dutch translation production task (L1 to L2 production), to tap participant's ability to retrieve the Dutch translations.[4]

---

[4] The English-to-Dutch production task was difficult for participants to perform because they were required to produce the Dutch words orally, although they were only provided a visual form during training. Given the relatively low accuracy on this task (less than 40%), the fact that providing the same translation twice was technically correct, and the fact that the patterns observed were largely similar to those found

In the free recall task, participants were asked to type all the Dutch words they remembered, along with the corresponding English translations they could recall. No time constraint was enforced for this task. For a response to be considered correct, participants had to provide the English word and its Dutch translation spelled correctly. In the Dutch-to-English translation production test, participants orally provided the English translations of the Dutch words, presented in a random order. On each trial, a fixation cross appeared until the participant pressed the space bar. A blank screen was then presented for 100 ms, followed by the Dutch word until the participant's vocal response triggered the voice key (E-Prime software, Psychology Software Tools, Pittsburgh, PA) or after 10 seconds, whichever came first. Participants' responses were digitally recorded for later coding of accuracy. In the versions administered in Sessions 1 and 2, only the 56 Dutch words that had been trained in that session were presented. In the version administered in Sessions 3 and 4, all 64 Dutch words that had been taught in that version of the study were presented.

In the translation recognition task, 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. Each trial began with a fixation cross until the participant pressed the space bar. Then, a blank screen appeared for 100 ms, followed by the Dutch–English word pair. The Dutch word appeared above the English word, on the same screen at the same time, until the participant made a response or after 5 seconds, whichever came first. This procedure differs slightly from studies that manipulated the presentation onset of the L1 vs. the L2 word (De Groot & Comijs, 1995), and is thus inappropriate for comparisons of translation recognition direction (L1 to L2 vs. L2 to L1). Sixty-four Dutch–English word pairs were presented in randomized order. Half of the pairs in each version were correct translations, and half were incorrect translations created by pairing Dutch words with the English translations of other Dutch words in the same condition.

To summarize, participants completed a total of six training cycles with 56 trials each, and were tested four times with an L2-to-L1 translation production task, twice with a free-recall task, once with a translation-recognition task, and once with an L1-to-L2 translation production task. They also completed three individual difference measures (Operation-Word span, Turner & Engle, 1989; Stroop, Stroop, 1935; color Simon, e.g., Bialystok, 2006) and completed a detailed language

history questionnaire (Tokowicz et al., 2004).[5] We focus our analysis and discussion on results from the tests administered in Sessions 3 and 4, because by that point in the experiment both translations in the separate condition had been taught. These tests reflect long-term retention, because no training took place during these sessions. For completeness, the results of the Dutch-to-English production tests administered at the end of Sessions 1 and 2 are presented in Table 6.

Results

Mean accuracy and reaction time by condition is provided in Table 6, and an overview of the results is given in Table 7. Note that we do not discuss the RT data in the text because these were largely not significant, and never contrasted with the pattern observed in the accuracy data. Instead, we focus our discussion on the accuracy data for which the findings were more pronounced (see also, Degani & Tokowicz, 2010). For the free recall and Dutch-to-English translation production tests, we examined the interactions between session (3 vs. 4) and ambiguity (translation ambiguous vs. unambiguous); session by ambiguity source (form-ambiguous vs. meaning-ambiguous); session by instructional method (together vs. separate); and session by presentation order (first vs. second in the separate condition). For the translation recognition test, administered only on Session 3, we examined the same effects without the by-session interactions. Figure 2 shows the instructional method effects across measures, and Figure 3 shows the presentation order effects across measures. Notably, the interactions of ambiguity source with instructional method or with presentation order did not reach significance in any of the accuracy analyses, and are therefore not reported in detail in the text. We return to this issue in the Discussion section.

Analyses by participants are reported as F_1 and analyses by items are reported as F_2. To allow comparisons across sessions, data from 28 participants who completed all sessions are reported.[6] In the free-recall tests, analyses by participants were performed on the percentage of the Dutch words within a given condition that were correctly recalled (and spelled) along with their English translation (of the total possible in that condition; i.e., 24 translation-unambiguous, 32 translation-ambiguous). Free-recall analyses by items were performed on the percentage of participants who correctly recalled each Dutch word and its English translation (of the total participants who saw that word in a given condition).

---

in the Dutch-to-English production task, we refrain from elaborating on this measure further.

[5] The discussion of individual differences is beyond the scope of the current paper.
[6] The same pattern of results was found in Session 3 when data from all 40 participants available for this session were included.

Table 6. *Mean performance across tests and sessions.*

		Condition			
Test and session (n)	Unambiguous	Ambiguous together	Ambiguous separate	Separate first	Separate second
		Accuracy as % correct (SD)			
Free recall 3 (28)	09 (09)[1]	15 (13)[2]	11 (1)[1,2*]	18 (19)[a]	05 (08)[b]
Free recall 4 (28)	14 (12)[1]	21 (17)[2]	16 (10)[1,2*]	23 (17)[a]	09 (10)[b]
Dutch-to-English production 1 (28)	58 (24)[1]	60 (22)[1]	NA	NA	NA
Dutch-to-English production 2 (28)	80 (19)[1]	76 (23)[1]	NA	NA	NA
Dutch-to-English production 3 (28)	63 (20)[1]	64 (23)[1]	50 (18)[2]	59 (23)[a]	41 (20)[b]
Dutch-to-English production 4 (28)	67 (18)[1]	66 (20)[1]	55 (16)[2]	61 (21)[a]	48 (20)[b]
Translation recognition 3 (28)-overall	95 (05)[1]	92 (07)[2]	88 (11)[2]	89 (14)[a]	87 (14)[a]
Yes responses	96 (06)[1]	94 (10)[1]	83 (17)[2]	84 (23)[a]	82 (23)[a]
No responses	95 (06)[1]	91 (11)[2]	95 (07)[1,2]	95 (13)[a]	95 (11)[a]
		Reaction times in ms (SD)			
Free recall 3 (28)	NA	NA	NA	NA	NA
Free recall 4 (28)	NA	NA	NA	NA	NA
Dutch-to-English production 1 (28)	1530 (468)[1]	1737 (589)[2]	NA	NA	NA
Dutch-to-English production 2 (28)	1222 (319)[1]	1315 (436)[2]	NA	NA	NA
Dutch-to-English production 3 (28)	1519 (471)[1]	1466 (436)[1]	1552 (506)[1]	1560 (563)[a]	1686 (613)[a]
Dutch-to-English production 4 (28)	1330 (358)[1,2]	1357 (428)[1]	1411 (353)[1,2*]	1234 (363)[a]	1660 (734)[b]
Translation recognition 3 (28)	1233 (312)[1]	1289 (330)[1*2*]	1345 (280)[2]	1316 (474)[a]	1437 (477)[a]

Note: Data are shown as a function of the type of item (unambiguous, ambiguous together, ambiguous separate, separate first vs. separate second), collapsing over ambiguity source (form vs. meaning). Means in the same row that do not share a numeric subscript differ at the $p < .05$ level, based on F_1 analyses comparing unambiguous to ambiguous items, with planned comparisons between the three types. Means in the same row that do not share an alphabetic subscript differ at the $p < .05$ level based on *t*-tests between items in the separate condition trained first vs. second. Standard deviations (SDs) are shown in parentheses. * indicates a marginal effect.

Table 7. *Results overview.*

		Effect									
		Session effect		Ambiguity effect		Ambiguity source		Instructional method		Presentation order	
Measure (Session)		F_1	F_2	F_1	F_2	F_1	F_2	F_1	F_2	F_1	F_2
Free recall (3&4)	Acc	**.000**	**.000**	**.023**	n.s.	n.s.	n.s.	**.026**	.055	**.000**	**.000**
Dutch-to-English production (3&4)	Acc	**.024**	n.s.	**.001**	n.s.	**.030**	n.s.	**.000**	**.000**	**.001**	**.000**
	RT	**.000**	**.000**	n.s.	n.s.	n.s.	n.s.	n.s.	**.025**	**.011**	**.002**
Translation recognition (3)	Acc	NA	NA	**.000**	**.001**	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
	RT	NA	NA	**.008**	.074	n.s.	n.s.	.080	n.s.	n.s.	n.s.

Note: Values reflect *p* values for significant or marginally **significant** effects based on analysis by participants (F_1) and by items (F_2) for accuracy (Acc) and reaction-time (RT) data in the different tests. In all relevant analyses, the interactions with session were n.s. in both F_1 and F_2.

Free recall

In the free-recall test, performance significantly increased from Session 3 to Session 4 (.11 vs. .16), F_1 (1, 27) = 24.06, $MSE = .003$, $p < .001$, $\eta_p^2 = .47$; F_2 (1, 70) = 27.19, $MSE = .003$, $p < .001$, $\eta_p^2 = .28$, but was still relatively low. Interestingly, an ambiguity ADVANTAGE emerged in the analysis by participants, F_1 (1, 27) = 5.79, $MSE = .008$, $p = .023$, $\eta_p^2 = .18$; F_2 (1, 70) = 2.35,

$MSE = .022$, $p = .13$, $\eta_p^2 = .03$, such that translation-ambiguous words were recalled more often ($M = .16$) than translation-unambiguous words ($M = .12$), and this effect did not vary by session, $Fs < 1$. The effect of ambiguity source was not significant, $Fs < 1$, and it did not vary by session, $Fs < 1$.

Instructional method significantly influenced free-recall performance, such that translation-ambiguous
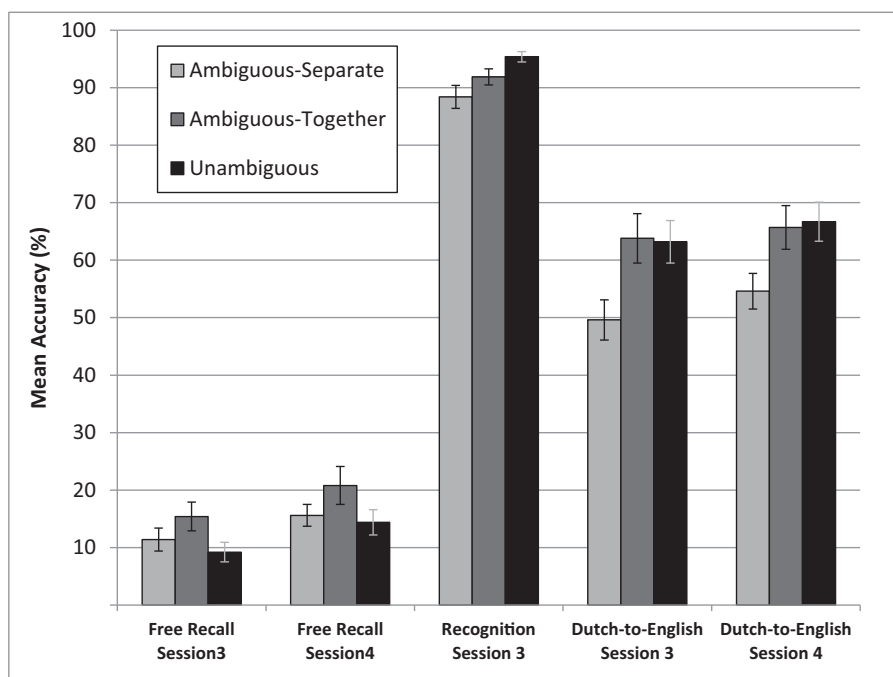
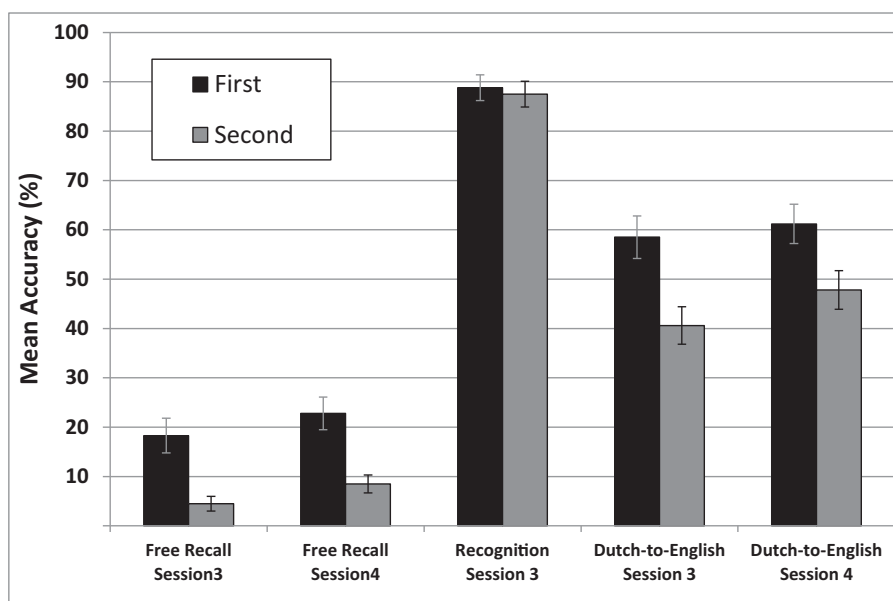Figure 2. Response accuracy as a function of instructional method in comparison to the unambiguous words.

Figure 3. Response accuracy as a function of presentation order for translation-ambiguous words taught in the separate condition.

words taught together were recalled more often ($M = .18$) than translation-ambiguous words taught separately ($M = .14$), $F_1 (1, 27) = 5.53$, $MSE = .011$, $p = .026$, $\eta_p^2 = .17$; $F_2 (1, 47) = 3.88$, $MSE = .023$, $p = .055$, $\eta_p^2 = .08$. The effect was stable across sessions, $Fs < 1$.

Within the translation-ambiguous words taught separately, those taught first were recalled significantly more often ($M = .20$) than those taught second ($M = .07$),

$F_1 (1, 27) = 17.78$, $MSE = .031$, $p < .001$, $\eta_p^2 = .40$; $F_2 (1, 47) = 24.49$, $MSE = .036$, $p < .001$, $\eta_p^2 = .34$. This effect was consistent across sessions, $Fs < 1$.

Dutch-to-English translation production

In the Dutch-to-English translation production test, performance significantly increased from Session 3 to

Session 4 (.60 vs. .63) in the analysis by participants, F_1 (1, 27) = 5.75, MSE = .006, p = .024, η_p^2 = .17; F_2 (1, 46) = 1.54, MSE = .019, p = .221, η_p^2 = .03. Importantly, an ambiguity DISADVANTAGE emerged in the analysis by participants, F_1 (1, 27) = 14.95, MSE = .008, p = .001, η_p^2 = .36; F_2 (1, 46) = 1.64, MSE = .065, p = .21, η_p^2 = .03, such that translation-unambiguous words were translated more accurately (M = .65) than translation-ambiguous words (M = .58), and this effect did not vary by session, Fs < 1.

According to the analysis by participants, meaning-ambiguous words were translated more accurately (M = .61) than form-ambiguous words (M = .56), F_1 (1, 27) = 5.22, MSE = .015, p = .030, η_p^2 = .16; F_2 < 1.

Translation-ambiguous words taught together were translated more accurately (M = .65) than translation-ambiguous words taught separately (M = .52), F_1 (1, 27) = 23.61, MSE = .019, p < .001, η_p^2 = .47; F_2 (1, 23) = 25.58, MSE = .016, p < .001, η_p^2 = .53. The effect was stable across sessions, F_1 (1, 27) = 1.12, MSE = .006, p = .299, η_p^2 = .04; F_2 < 1. Within the translation-ambiguous words taught separately, those taught first were translated significantly more accurately (M = .60) than those taught second (M = .44), F_1 (1, 27) = 14.63, MSE = .047, p = .001, η_p^2 = .35; F_2 (1, 23) .034, p < .001, η_p^2 = .43.

Translation recognition

Translation recognition data were available only for Session 3. On this test, translation-unambiguous words were recognized more accurately (M = .95) than translation-ambiguous words (M = .90), F_1 (1, 27) = 16.73, MSE = .002, p < .001, η_p^2 = .38; F_2 (1, 46) = 12.56, MSE = .003, p = .001, η_p^2 = .22.[7] The effect of ambiguity source was not significant, F < 1. Instructional method did not significantly affect translation recognition accuracy, F_1 (1, 27) = 2.60, MSE = .007, p = .118, η_p^2 = .09; F_2 (1, 23) = 1.24, MSE = .010, p = .276, η_p^2 = .05, and similarly, presentation order within the separate condition did not significantly affect performance, Fs < 1.

Discussion

In accordance with previous work (Degani & Tokowicz, 2010), the results of the current study demonstrate a difficulty in learning translation-ambiguous words compared to translation-unambiguous words, in translation production and recognition tests across various

time points. Extending previous research, however, the current study shows that this learning disadvantage does not generalize across all tasks or items, and further that it can be alleviated under some instructional conditions. In particular, the results show that translation-ambiguous words taught together, on consecutive trials in the same training session, are learned better than the same translation-ambiguous words taught in separate training sessions. This instructional effect was prominent in participants' accuracy, in both free recall and Dutch-to-English production tasks. This pattern of findings suggests that this instructional method improves the learner's ability to accurately retrieve the meaning and L1 translation of the newly-acquired Dutch word.

The advantage in the current study for items presented together over items presented separately aligns with other findings in the literature. In particular, McDaniel et al. (2001) presented four words associated with a single icon (similar to translation-ambiguous words) on an AAC device together (Experiment 1; similar to the together condition) or interspersed with other items (Experiment 2, blocked by word type rather than by icon; similar to the separate condition). Comparison of the means across the two experiments suggested an advantage for blocked presentation. The current results provide further support for such an advantage when items have a many-to-one mapping.

Presenting translation-ambiguous words together may guide learner's attention toward distinguishing attributes of the two translations (Underwood, 1969; see also Postman & Gray, 1977), allowing beginning L2 learners to maintain a one-to-one mapping more easily, when possible. In addition, the advantage of teaching translation-ambiguous words together may be rooted in the *timing* with which information was presented. In particular, presenting learners with both translations from the beginning may allow learners to create the appropriate one-to-many mapping structure, avoiding the need to later revise the mapping they created. Thus, allocating the appropriate association strength to each possible translation from the beginning appears to improve learning in comparison to a condition in which the inappropriate weight is given to a single translation, which later has to be adjusted and divided with an additional correct translation. The results of Medina et al. (2011) similarly suggest that starting out correctly leads to superior performance, because learners achieved better ultimate performance when more informative cues were presented to them early during training. The authors suggested that learners generate a single conjecture about a word and seek its confirmation on later encounters. Revising it when it turns out to be erroneous requires more evidence and effort.

The current results further suggest that the need to create the appropriate one-to-many mapping structure is

---

[7] The translation-ambiguity disadvantage was similarly evident in a d' analysis that adjusted for response biases, F_1 (1, 39) = 30.62, MSE = 1.33, p < .001, η_p^2 = .44. Translation-ambiguous words were associated with lower d' than translation-unambiguous words (3.23 vs. 4.65).

item specific and does not reflect a *general* awareness of the presence of translation ambiguity. Indeed, all participants encountered translation ambiguity from the beginning, not only because they were informed that some English words would have more than one Dutch translation, but also because they were presented with some translation-ambiguous words in the together condition (i.e., the two translations were presented on consecutive trials in both training sessions). Apparently this general awareness was not sufficient to prepare the learner for translation ambiguity that surfaced in later sessions. Presumably, it is difficult to create a flexible mapping structure that would become useful for both ambiguous and unambiguous words. Thus, when presented with only one translation for a particular English word in Session 1, learners seem to have created a one-to-one mapping that they later had to revise, creating difficulty in learning items in the separate condition.

When the ambiguity of a given item is presented from the beginning in the together condition, learners appear able to handle this one-to-many mapping substantially better than when the ambiguity surfaces only later. In fact, in the Dutch-to-English translation production tests, teaching both translations together led to performance that was statistically indistinguishable from performance on translation-unambiguous words (see Table 6). Thus, providing learners with relevant information earlier during training aids the learner considerably. This conclusion is further supported by the current findings regarding the order of presentation in the separate condition. Specifically, the results demonstrate a strong and reliable advantage for the translations learned in Session 1 over those learned in Session 2. This advantage was prominent in the accuracy performance on the recall and production tests. This finding extends previous work demonstrating an order-of-acquisition effect on speed of processing in foreign vocabulary learning (Izura et al., 2011).

Further, the results of the current study show that an advantage for the item learned first (e.g., Poepsel et al., 2012) is not limited to cases in which that item is not related to later items (as in Izura et al., 2011), but holds for associated representations. This type of competition resembles classical paired-associate paradigms in which second-list associations (A–C) are related to first-list associations (A–B) through the source word (A). Tulving and Watkins (1974) observed an advantage for first- over second-list associations when no first-list associations were tested immediately before second-list learning. In the current study, we did not present or test learners with the first translation learned in Session 1 before presenting the second translation in Session 2 in the separate condition, but we did present learners with other items from Session 1. Our results nonetheless demonstrate a strong order/priority effect for translations learned first.

The observed advantage for translation-ambiguous items taught together versus separately could be explained by the difference between these two types of items in the number of tests administered during training and by the difference in spacing throughout training (e.g., Rawson & Dunlosky, 2011). In particular, because items in the together condition were trained three times in Session 1 and three times in Session 2, they were tested in a Dutch-to-English production test twice before the reported test administered in Session 3. In contrast, items in the separate condition were taught six times either in Session 1 or Session 2, and were thus tested only once before Session 3. Previous research suggests that testing is an efficient learning technique (e.g., McDaniel, Anderson, Derbish & Morrisette, 2007). To tease apart the contribution of these two sources (testing effect vs. instructional method) we conducted an additional analysis in which we compared performance in the together condition from Session 3 to performance in the separate condition from Session 4. At these time points, both types of items had been taught six times and tested twice previously. Items taught together still exhibited significantly better performance than items taught separately (64 vs. 55%), despite an overall significant improvement in performance from Session 3 to Session 4 in both accuracy and reaction time. Thus, the observed advantage for teaching translation-ambiguous items together cannot be reduced to the number of tests incorporated during training.

With respect to spacing of training cycles, teaching of items in the together condition was interleaved with a two-day break (between Sessions 1 and 2), whereas teaching of items in the separate condition was blocked in one session. The English words themselves, however, were equally spaced in that they were presented six times in Session 1 and six times in Session 2. Nonetheless, it remains to be tested whether the advantage for teaching together vs. separately holds when spacing is equated for by the foreign-language word.

Notably, the confound between instructional method and testing and spacing is likely a reflection of real-life foreign-vocabulary learning, in which words learned earlier have more opportunities to be tested, and their teaching may be spaced throughout the semester. Indeed, if instructors opt to present the two alternative translations of ambiguous words from the beginning, rather than withholding one translation until later weeks or semesters, both translations would benefit from more testing and could benefit from interleaved teaching.

Four results from the current investigation deserve special attention. First, the effects in the current study are manifested mostly in participants' accuracy (see also Degani & Tokowicz, 2010), likely because our tests tap earlier points on the learning curve (see also e.g., DeKeyser & Larson-Hall, 2005), at which time the representations are less stable. Indeed, when accuracy is

low, fewer trials are available for reaction-time estimation, leading to less stable estimates.

Second, we found that, overall, translation-ambiguous words were recalled *better* than translation-unambiguous words in a free-recall test, both one and three weeks after initial learning. This is in contrast to the disadvantage of these translation-ambiguous words in production and recognition tests (see also Degani & Tokowicz, 2010). This result may suggest that translation-ambiguous words were linked to more representations during learning (such as the other Dutch translation), allowing better performance on an unstructured free-recall test. In this type of test, participants have more time to resolve any competition that may arise from the alternative translation. Alternatively, it may be that translation-ambiguous words, like lower-frequency words, tend to attract more attention during learning of mixed lists, therefore leading to better recall (e.g., DeLosh & McDaniel, 1996). Further, because participants were aware that some words were translation-ambiguous, as discussed above, they may have invested more effort and resources in memorizing such words, or may have prioritized them in the active free-recall task. In the more structured cued-recall tests, such bias was less prominent. Indeed, the advantage for translation-ambiguous words in the free-recall task appears to be driven by the items taught together (see Figure 2), for which the two Dutch translations can easily serve as effective cues for each other. However, it is important to note that performance on the free-recall tests was still extremely low (below 21%). In addition, due to the unconstrained nature of the task, when incorrect translation pairs were retrieved, it was not possible to determine whether the Dutch word was accurately retrieved but its meaning in English was incorrectly retrieved, or whether participants accurately retrieved the English word but paired it with an incorrect Dutch translation. The translation-ambiguity advantage in an unconstrained free recall task awaits further replication.

Third, we found that the effect of instructional method operated similarly for form- and meaning-ambiguous items. As alluded to earlier, form-ambiguity encompasses one-to-many mappings at both the lexical and conceptual level (see Figure 1), whereas the one-to-many mapping could be limited to the lexical level for meaning-ambiguous items. The fact that we observed similar instructional-method effects for both types of items suggests that manipulating the presentation of the two alternatives (together vs. separately) operated mainly at the lexical level. Specifically, presenting both alternatives together aided participants in assigning the appropriate associative strength to each of the two lexical links, and in so doing prevented them from needing to revise their initial association structure. It is also possible that presenting them together allowed learners to become aware of distinctive features at the lexical level

(phonological and orthographic cues). Learners at these early stages of learning may be more tuned to lexical rather than semantic aspects of the to-be-learned word (Kroll & Stewart, 1994), although there is some evidence to suggest that this may not necessarily be the case (Brysbaert & Duyck, 2010; Degani & Tokowicz, 2010).

Fourth, performance improved overall from Session 3 to Session 4. This is in contrast to the results of Degani and Tokowicz (2010), which showed a decline in performance across sessions. Note, however, that here we do not compare immediate learning (Session 1 and 2) to long-term retention (Session 3), as was done in the earlier study, but rather compare performance on a long-term test (Session 3, one week after learning) to an even longer-term test (Session 4, three weeks after learning). The slight improvement in performance may be due to better familiarity with the tasks on the later session, or to the positive effect of testing on learning (e.g., Karpicke & Roediger, 2007). The stability of the critical results (ambiguity, instruction, and order effects) across sessions highlights the enduring nature of the findings.

Together, these findings point to the important role of order of acquisition in foreign-language vocabulary learning. Setting the appropriate one-to-many mapping structure, or dividing up the association strength among the relevant candidates, from the beginning seems to aid learners in acquiring translation-ambiguous words. In addition, the order effect shows that the translation learned first enjoys a continued advantage over that learned second, pointing to an important finding for education.

To the extent that one nonetheless resorts to the separate presentation schema, it is important to remember that in many cases one of the translations is more dominant than the other, in that it is more frequently contextually-appropriate (Eddington & Tokowicz, 2013; Laxén & Lavaur, 2010; Prior, Kroll & MacWhinney, 2013). In the current study, we treated both alternative translations as equally appropriate and taught them with equal frequency. However, given that in natural use typically one would be more dominant, it may be advantageous to teach it first, so as to provide it the appropriate advantage in use. Alternatively, it may be the case that the more dominant translation requires no additional boost, and it would be advantageous to teach the subordinate translation first. This issue remains to be examined in future investigations.

Finally, because translation ambiguity is widespread (e.g., Prior et al., 2007; Tokowicz et al., 2002) learners are likely to encounter many foreign-language vocabulary words that map in a many-to-one fashion to their L1 translation or meaning. To the extent that learners are ultimately to acquire both translations, it is advantageous to teach both together. The current results show that withholding the second translation leads to poorer learning, especially of the second translation.

Appendix. Complete set of stimuli and definitions by condition

Condition	English word	Definition	Dutch translation(s)
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
Meaning-ambiguous	block	1. a three-dimensional shape with six square or rectangular sides	1. blok
		2. to prevent access or progress	2. blokkade
Meaning-ambiguous	proposal	1. something offered or suggested, such as a plan	1. voorstel
		2. an offer of marriage	2. aanzoek
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

Appendix. Continued

Condition	English word	Definition	Dutch translation(s)
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
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.

Anderson, M. C., & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (eds.), *Handbook of perception and cognition: Memory* (2nd edn.), pp. 237–313. San Diego, CA: Academic Press.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39,* 445–459.

Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Canadian Journal of Experimental Psychology, 60,* 68–79.

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

De Groot, A. M. B., & Comijs, H. (1995). Translation recognition and translation production: Comparing a new and an old tool in the study of bilingualism. *Language Learning, 45,* 467–509.

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

De Groot, A. M. B., & Van Hell, J. G. (2005). The learning of foreign language vocabulary. In Kroll & De Groot (eds.), pp. 9–29.

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

DeKeyser, R., & Larson-Hall, J. (2005). What does the critical period really mean? In Kroll & De Groot (eds.), pp. 88–108.

DeLosh, E. L., & McDaniel, M. A. (1996). The role of order information in free recall: Application to the word-frequency effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1136–1146.

Eddington, C. M., & Tokowicz, N. (2013). Examining English–German translation ambiguity using primed translation recognition. *Bilingualism: Language and Cognition, 16,* 442–457.

Izura, C., Pérez, M. A., Agallou, E., Wright, V. C., Marín, J., Stadthagen-González, H., & Ellis, A. W. (2011). Age/order of acquisition effects and the cumulative learning of foreign words: A word training study. *Journal of Memory and Language, 64,* 32–58.

Karpicke, J. D., & Roediger, H. L. III (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language, 57,* 151–162.

Karpicke, J. D., & Roediger, H. L. III (2008). The critical importance of retrieval for learning. *Science, 319,* 966–968.

Kaushanskaya, M., & Marian, V. (2009). Bilingualism reduced native-language interference during novel word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 829–835.

Kroll, J. F., & De Groot, A. M. B. (eds.) (2005). *Handbook of bilingualism: Psycholinguistic approaches*. New York: Oxford University Press.

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.

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.

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

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.

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.

McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology, 19,* 494–513.

McDaniel, M. A., Neufeld, K. H., & Damico-Nettleton, S. (2001). Many-to-one and one-to-many associative learning in a naturalistic task. *Journal of Experimental Psychology: Applied, 7,* 182–194.

Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences, 108,* 9014–9019.

Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning, 60,* 154–193.

Pasquarella, A., Gottardo, A., & Grant, A. (2012). Comparing factors related to reading comprehension in adolescents who speak English as a first (L1) or second (L2) language. *Scientific Studies of Reading, 16,* 475–503.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11,* 357–383.

Poepsel, T., Gerfen, C., & Weiss, D. J. (2012). Context, mutual exclusivity, and the challenge of multiple mappings in word learning. *Proceedings of the 36th Annual Boston University Conference on Language Development* (vol. 2), pp. 474–486.

Postman, L., & Gray, W. (1977). Maintenance of prior associations and proactive inhibition. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 255–263.

Postman, L., & Parker, J. F. (1970). Maintenance of first-list associations during transfer. *The American Journal of Psychology, 83,* 171–188.

Postman, L., & Underwood, B. J. (1973). Critical issues in interference theory. *Memory and Cognition, 1,* 19–40.

Potts, R., & Shanks, D. R. (2012). Can testing immunize memories against interference? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 1780–1785.

Prior, A., Kroll, J. F., & MacWhinney, B. (2013). Translation ambiguity but not word class predicts translation performance. *Bilingualism: Language and Cognition, 16,* 458–474

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.

Rawson, K. A., & Dunlosky, J. (2011). Optimizing schedules of retrieval practice for durable and efficient learning: How much is enough? *Journal of Experimental Psychology: General, 140,* 283–302.

Stroop, J. R. (1935). Studies in interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.

Tokowicz, N., & Degani, T. (2010). Translation ambiguity: Consequences for learning and processing. In B. Van

Patten & J. Jegerski (eds.), *Research on second language processing and parsing*, pp. 281–293. Amsterdam: John Benjamins.

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, 34,* 435–451.

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.

Torres, I. J., Flashman, L. A., O'Leary, D. S., & Andreasen, N. C. (2001). Effects of retroactive and proactive interference on word list recall in schizophrenia. *Journal of the International Neuropsychological Society, 7,* 481–490.

Tulving, E., & Watkins, M. J. (1974). On negative transfer: Effects of testing one list on the recall of another. *Journal of Verbal Learning and Verbal Behavior, 13,* 181–193.

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

Underwood, B. J. (1969). Attributes of memory. *Psychological Review, 76,* 559–573.

Wahlheim, C. N., & Jacoby, L. L. (2011). Experience with proactive interference diminishes its effects: Mechanisms of change. *Memory and Cognition, 39,* 185–195.

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