



The role of polysemy in masked semantic and translation priming

Matthew Finkbeiner,^{a,*} Kenneth Forster,^b Janet Nicol,^b
and Kumiko Nakamura^b

^a *Department of Psychology, Harvard University, Cambridge, MA 02138, USA*

^b *Department of Psychology, University of Arizona, Tucson, AZ 85721, USA*

Received 1 December 2003; revision received 5 January 2004

Available online 5 March 2004

Abstract

A well-known asymmetry exists in the bilingual masked priming literature in which lexical decision is used: namely, masked primes in the dominant language (L1) facilitate decision times on targets in the less dominant language (L2), but not vice versa. In semantic categorization, on the other hand, priming is symmetrical. In Experiments 1–3 we confirm this task difference, finding robust masked L2–L1 translation priming in semantic categorization but not lexical decision. In formulating an account for these findings, we begin with the assumption of a representational asymmetry between L1 and L2 lexical semantic representations, such that L1 representations are richly populated and L2 representations are not. According to this representational account, L2–L1 priming does not occur in lexical decision because an insufficient proportion of the L1 lexical semantic representation is activated by the L2 prime. In semantic categorization, we argue that the semantic information recruited to generate a decision is restricted by the task category, and that this restriction enhances the effectiveness of the L2 prime. In Experiments 4–6, these assumptions were tested in a within-language setting by pairing many-sense words (e.g., “head”) with few-sense words (e.g., “skull”). In lexical decision, robust priming was obtained in the many-to-few direction (analogous to L1–L2), but, no priming was obtained in the few-to-many direction (analogous to L2–L1) using the same word pairs. Priming in semantic categorization, on the other hand, was obtained in both directions. We propose the Sense Model as a possible account of these findings.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Masked translation priming; Semantic priming; Bilingual lexical processing; Bilingual lexicon; Lexical decision; Semantic categorization

The masked translation priming paradigm has proven to be instrumental in the development of psycholinguistic accounts of bilingual lexical representation and processing. One important finding that this paradigm has revealed, and one that any successful model of bilingual lexical processing must be able to explain, is the translation priming asymmetry in lexical decision

(Is it a word?). This well-established asymmetry, which has been reported by several researchers working with different bilingual populations, is characterized by the following pattern of findings: masked (subliminal) primes in the dominant language (L1) facilitate recognition of translation-equivalent words in the non-dominant language (L2) (de Groot & Nas, 1991; Gollan, Forster, & Frost, 1997; Jiang, 1999; Jiang & Forster, 2001; Keatly, Spinks, & De Gelder, 1994; Williams, 1994); however, masked L2 primes do not reliably facilitate recognition of L1 translation-equivalent

* Corresponding author. Fax: 1-617-496-6262.

E-mail address: msf@wjh.harvard.edu (M. Finkbeiner).

words (Gollan et al., 1997; Grainger & Frenck-Mestre, 1998; Jiang, 1999; Jiang & Forster, 2001; Keatly et al., 1994; Sanchez-Casas, Davis, & Garcia-Albea, 1992).¹ As we will see below, most accounts of the translation priming asymmetry have appealed to a “limiting factor” explanation, whereby a property specific to the L2 lexicon is thought to limit priming in the L2–L1 direction.

One obvious possibility is that words in L2 are not processed effectively when they are masked; however, the same subjects exhibit robust within-L2 masked repetition priming effects (Finkbeiner, in press; Gollan et al., 1997; Jiang, 1999; Jiang & Forster, 2001). That is, bilinguals are faster to respond to a target in their L2 (e.g., “HOUSE”) when it is preceded by a masked presentation of the same word but in different case (e.g., “house”) than they are when the target is preceded by a control prime (e.g., “truck”). This finding suggests that a lack of L2–L1 translation priming cannot be due to an inability to effectively process the masked primes.

The Revised Hierarchical Model (Kroll & Stewart, 1994; Kroll & Tokowicz, 2001), a dominant model in the field, accounts for the translation priming asymmetry by suggesting that: (1) the locus of the translation priming effect is at the level of meaning and that (2) relative to L1 representations, L2 lexical forms are only weakly connected to meaning-level representations. According to this account, priming is effective in the L1–L2 direction because the masked L1 prime serves to activate a shared conceptual node, which then preactivates the L2 translation-equivalent lexical form. However, priming is not effective in the L2–L1 direction because L2 primes do not automatically activate their conceptual representations, resulting in no preactivation of the L1 translation-equivalent form and thus no priming. According to this account, the L2 form–meaning connection strength is the limiting factor that prevents priming in the L2–L1 direction.

This account is appealing, but it suffers from a major weakness in that it predicts no within-L2 priming. If the locus of priming is at the level of meaning, and L2 forms cannot automatically activate their meanings, then how does masked L2–L2 priming occur? To explain the existence of within-L2 priming, a second locus of priming would have to be posited: namely, it would have to be argued that within-L2 priming occurs at the level of form. In support of this claim, it could be pointed out that within-L2 priming effects have primarily been demonstrated with repetition priming, not semantic priming (Gollan et al., 1997; Jiang, 1999; although see

Frenck-Mestre & Prince, 1997), and consequently may be attributed to priming at the form level alone. However, translation priming effects with cognate stimuli (e.g., *rico-RICH*) cannot be attributed to an overlap in form-level properties alone, since the priming is measured relative to a baseline involving a similar degree of form overlap (e.g., *riect-RICH*). Therefore it appears that cognate priming must be attributed to an overlap of information at the level of meaning.

A related problem for the RHM is its underspecification with respect to the nature of the L2 form–meaning connection. This is presented most clearly in the awkward proposal that the L2 form–meaning connection can simultaneously support robust priming in the L1–L2 direction and limit priming in the L2–L1 direction. To explain both of these findings, it is assumed that the mapping between L2 form and meaning is sufficiently strong in the meaning-to-form direction (thereby permitting L1–L2 priming), but that the same mapping is too weak in the form-to-meaning direction (thereby limiting L2–L1 priming).

Recently, another finding using the masked translation priming paradigm has been reported which promises to elucidate the relationship between L1 and L2 lexical representations. This is the possibility of a task difference in masked L2–L1 repetition priming. To date, there have been two reported instances of such a task difference. The first was reported by Grainger and Frenck-Mestre (1998), who observed masked translation priming effects in the L2–L1 direction for non-cognates when semantic categorization was used, but, in line with several other researchers, not when lexical decision was used. This finding appears to be quite inconsistent with the notion that L2 words are incapable of activating semantic representations when masked, and hence it is critical to examine this issue further. This is especially true since the Grainger and Frenck-Mestre (1998) finding stands in marked contrast to that of Sanchez-Casas et al. (1992), who also used the semantic categorization task, but reported no masked L2–L1 translation priming unless the translation pairs were cognates.

The second example of a task effect was reported by Jiang and Forster (2001), who observed masked L2–L1 translation priming when an “old–new” episodic recognition task was used, but, again, not when lexical decision was used. Their explanation of this effect, which we refer to as the “separate memory systems account,” offers a very different interpretation of bilingual lexical processing. In their study, participants were asked to memorize a list of L1 words in the first phase of the experiment. In the second phase, participants performed a speeded “old–new” task, in which L1 words had to be classified according to whether they were on the original list or not. Unknown to the participants, the L1 target word was preceded by a masked translation prime in L2.

¹ The translation priming asymmetry is eliminated when prime and target words are cognates from same-script languages (e.g., *rico-RICH*) (Sanchez-Casas et al., 1992).

Using this procedure, Jiang and Forster (2001) obtained significant masked translation priming effects in the L2–L1 direction for “old,” but not “new” targets. Using the same materials and two-phase design, no priming was obtained when the task was changed to lexical decision. Jiang and Forster explained their results by proposing that L1 and L2 lexicons are stored in separate memory systems: the L1 lexicon is stored in lexical memory, while the L2 lexicon is stored in episodic memory as a set of associations between L2 words and their L1 counterparts (represented episodically). If responses are controlled by activation patterns in episodic memory (as they would be in an “old–new” task), then L2–L1 priming can easily be explained. In contrast, when responses are controlled by activation patterns in lexical memory (as they would be in a lexical decision or semantic categorization task), the L2 prime is thought to be unable to activate the lexical representation of its L1 counterpart. This is due to a critical assumption of the separate memory systems account, which holds that activation across memory systems requires awareness. Although this account can explain the task difference between lexical decision and episodic recognition, it cannot straightforwardly explain the results obtained when semantic categorization is used. Since semantic categorization clearly requires that the lexical representation of the L1 target be activated, it should yield the same effects as the lexical decision task, i.e., no L2–L1 priming. Hence, a task difference where masked translation priming is observed in semantic categorization but not lexical decision would be a challenge to both the connection-strength explanation provided by the RHM as well as the explanation laid out in the separate memory systems account.

The purpose of the present study is: (1) to confirm the task difference originally reported by Grainger and Frenck-Mestre (1998) between semantic categorization and lexical decision in masked L2–L1 translation priming and (2) to propose and test a theoretical account of masked translation priming that at once is able to accommodate the task difference, the masked translation priming asymmetry, and the robust within-L2 masked priming effects. To anticipate our results, we found masked translation priming in the L2–L1 direction using semantic categorization but not lexical decision. In the second part of the article, we propose an account of these findings, which we tentatively refer to as the “Sense Model.” The Sense Model is different from other models of bilingual lexical processing in that it proposes that priming between semantically related words depends on the *proportion* of shared senses. It seems reasonable to suppose that L1 words are associated with many more semantic senses than their L2 counterparts. From this it follows that the proportion of primed senses belonging to the target word will be much higher when an L2 word is primed by its L1

counterpart than when an L1 word is primed by its L2 counterpart. Note that similar representational asymmetries exist between semantically related words within a single language too (e.g., “head” has many senses where “skull” only has one or two). Hence, it should be the case that priming asymmetries exist for within-language priming as well (e.g., “head” should prime “skull,” but “skull” should not prime “head”). In the third part of the article, we test specific predictions of the Sense Model, using within-English semantically related prime–target word pairs. To anticipate our results once again, we found that we were able to recreate the priming asymmetry seen in translation priming when the selection of materials adhered to the assumptions of the Sense Model for translation equivalents. That is, in lexical decision, priming occurred between many- and few-sense words in the many-to-few (analogous to “L1–L2”) direction, but not in the few-to-many (or “L2–L1”) direction. Furthermore, we observed priming in both directions in the semantic categorization task, thereby recreating the (cross-language) task difference between lexical decision and semantic categorization. Taken together, these findings provide strong support for the assumptions of the Sense Model. More generally, these findings present difficult challenges to other current models of bilingual lexical representation and processing.

Experiment 1—replication of Grainger and Frenck-Mestre (1998)

A comparison of the Sanchez-Casas et al. (1992) study and the Grainger and Frenck-Mestre (1998) study reveals several differences that could have produced the contradictory findings. Perhaps the most striking difference is in how the categories were presented to the participants. It is often the case in semantic categorization tasks that items are presented in a blocked design such that all of the exemplars (and an equal number of non-exemplars) appear together (Bueno & Frenck-Mestre, 2002; Forster & Hector, 2002; Forster, Mohan, & Hector, 2003; Frenck-Mestre & Bueno, 1999). Grainger and Frenck-Mestre (1998) used this procedure, whereas in the Sanchez-Casas study, categories changed on each successive item. A concern with a category switching procedure is that each experimental item may effectively constitute a separate task, and the consequence of this “task-switching paradigm” may be diminished effects. As an example in support of such a possibility, Catchpole (1987) found that blocking items into categories restored the frequency effect for non-exemplars, an effect that had previously gone undetected in experiments using different categories on each trial (Balota & Chumbley, 1984). If switching categories on each successive trial can serve to eliminate an effect as

robust as the frequency effect, then it follows that blocking items according to semantic category in semantic categorization tasks may be necessary to observe any lexical processing effects, including translation priming effects. Accordingly, items in the present experiment were blocked according to category, and participants were given practice items with each new category.

Just as there is reason to be concerned that the design of the Sanchez-Casas et al. study may have led to a Type II error, there is also reason to be concerned about particular aspects of the Grainger and Frenck-Mestre study. For example, there were a relatively small number of participants in the Grainger and Frenck-Mestre study ($N = 12$; two of whom were the authors, as opposed to an N of 21 in the Sanchez-Casas study) and each item was presented a total of 24 different times to each participant. Obviously, this degree of repetition may lead to strategic effects. Similarly, the dominant language of the English–French bilinguals in the Grainger and Frenck-Mestre study may have been their L2 French (as a result of living in France for many years), whereas the participants in the Sanchez-Casas et al. study were certainly not dominant in their L2 English. If this were the case, it may explain why Grainger and Frenck-Mestre observed priming in the French–English (L2–L1) direction and why Sanchez-Casas et al. did not (except for cognates) in the English–Spanish (again L2–L1) direction. In the present experiment, the dominant language of all of our Japanese–English bilinguals was their L1 (Japanese).

Method

Participants

Twenty Japanese–English bilinguals were recruited from the University of Arizona campus community. All participants were native speakers of Japanese and were employed as graduate students at the University of Arizona. Participants had received a minimum of 6 years of English instruction while in Japan and at the time of testing all had been living in the United States for at least 2 years. All participants were paid for their participation.

Materials

In order to ensure translation equivalency for each English–Japanese prime–target word pair, 5 Japanese–English bilinguals (from the same population as the participants in the experiment) were asked to translate a list of 170 items from English into Japanese (L2–L1); another group of 5 was asked to translate the same items in the opposite direction (L1–L2). Only those word pairs that were translated identically in each direction by all participants were selected as critical items. Fifty-two word pairs met this criterion. The

critical items belonged to 11 different semantic categories (*part of building, family relative, color, unit of time, animal, profession, insect, scientific discipline, part of the body, kind of metal, reading material*). In order to make use of a blocked design, an additional 58 items were selected to serve as “practice” items on trials preceding the critical items in each category. This resulted in a total of 110 “Yes” items (10 per category), with a minimum of 4 per category being critical. Additionally, 110 non-exemplar targets (NO items) were chosen. These were chosen so as to ensure that they could not be construed as belonging to any of the 11 categories. An additional 220 English words were selected to serve as primes on control and practice exemplar trials, as well as on non-exemplar trials. These were unrelated to their targets, but were matched with the critical primes for frequency, concreteness, imageability, and word-length.

Care was taken to ensure that none of the Japanese targets shared cognate status with their English primes. All targets were presented in their standard script, which was Kanji for all categories except for those appearing in the INSECT category (the standard script for insects is Katakana).

Design and procedure

Items were blocked according to semantic category. For each category, there were 10 exemplar and 10 non-exemplar trials. Each trial consisted of the following sequence (adapted from Grainger and Frenck-Mestre): first, the participant was presented with a forward mask (#####) for 500 ms, followed by an English prime (translation or control) in lowercase letters for 50 ms, followed by a backward mask for 150 ms, and then the Japanese target word for 500 ms. The backward mask was interpolated between the prime and the target in order to ensure that participants had sufficient time to process the L2 prime. It has been suggested that L2–L1 priming may not occur because of a speed of processing asymmetry (cf. Jiang, 1999). According to this argument, the processing of L2 primes lags too far behind the processing of L1 targets to produce any observable priming. Considering this possibility, we presented a backward mask for a relatively long duration (150 ms for a total SOA of 200 ms). A potential concern with this procedure is that the prime will appear to “pop out” of the background created by identical forward and backward masks. In order to prevent this possibility, the backward mask differed from the forward mask both in font type (Arial Black, as opposed to Times New Roman) and size (18 as opposed to 11). This particular presentation sequence was apparently effective in preventing participants from becoming aware of the presence of the prime as they all expressed surprise during their debriefing to learn that an English word had preceded the Japanese target.

Half of the critical targets per list were preceded by their translation equivalents and half were preceded by a control prime. No other targets were preceded by translation-equivalent primes. Two counterbalanced lists were constructed such that if a target was preceded by its translation prime on List A, it was preceded by its control prime on List B and vice versa. No target word or prime word was repeated within lists. Participants were asked to indicate whether the target belonged to the category by pressing either a YES button or a NO button. Stimuli were presented randomly within categories, using the DMDX package developed at the University of Arizona by J.C. Forster (Forster & Forster, 2003), with the only constraint that the “practice” exemplars preceded the critical items. Response times (RTs) were recorded to the nearest millisecond.

Results and discussion

Data from trials on which an error occurred were discarded and outliers were replaced with values equal to cutoffs established 2 *SD* units above and below the mean for each participant. Mean response times were 475 ms in the translation prime condition, and 494 ms in the control prime condition. An ANOVA showed that this 19 ms priming effect was highly significant ($F(1, 19) = 12.84$, $P = .002$; $F(1, 51) = 24.38$, $P < .001$). The mean error rate was 1.2% and did not differ between conditions.

The findings of Experiment 1 reveal that our participants possessed sufficient L2 lexical processing skills to exhibit reliable masked translation priming effects in the L2–L1 direction in a semantic categorization task. As such, these results confirm the findings of Grainger and Frenck-Mestre (1998).² As we pointed out in the Introduction, it is not clear how the separate memory systems account could accommodate these findings. As for the connection-strengths account of the translation priming asymmetry, these findings indicate that, for these particular bilinguals, L2 form–meaning connection strengths did not constitute a limiting factor in L2–L1 priming. The connection-strength hypothesis should therefore predict that these particular bilinguals should also exhibit L2–L1 priming when the task is changed to lexical decision. We test this possibility in Experiment 2, using the same materials and masking procedure, changing only the task to lexical decision.

² This finding also suggests that the most effective procedure is to block items by category, rather than changing category from trial to trial, as was the case in the Sanchez-Casas et al. (1992) study.

Experiment 2—masked priming with L2 primes in lexical decision

Method

Participants

Eighteen Japanese–English bilinguals were recruited for the lexical decision experiment. Fourteen of these 18 participants had participated in the semantic categorization experiment 6 months prior.

Materials

The materials for the L2–L1 translation priming condition were identical to those used in the semantic categorization task except for the 110 Japanese non-words that had to be created. These were created by dividing a column of two-character Japanese words into two columns and randomly sorting one of those columns and then rejoining the two columns. Any resulting combinations that were thought by the fourth author (a native speaker of Japanese) to be words were discarded. Additionally, 80 English words were selected for the within-L2 repetition priming condition. Half of these were abstract words and half were concrete words. Concreteness values were taken from the MRC psycholinguistic database (available on the web at http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm). The abstract words had a mean concreteness value of 244.6 (on a scale of 100–700); the concrete words had a mean concreteness value of 592.6. The words were further divided on the basis of frequency. Half of the words were high-frequency items (mean frequency of 279 occurrences per million; CELEX) and half were low-frequency items (mean frequency of 5.73 occurrences per million). The purpose of selecting these particular stimuli was to see how closely the lexical performance of these bilinguals patterned to that of native speakers. If the bilingual participants exhibit clear frequency and concreteness effects, we can be relatively sure that they are processing these items in largely the same way as native speakers.

In addition, 80 non-word targets were selected to elicit “No” responses. Non-words were generated by the ARC Nonword Database (<http://www.macqs.mq.edu.au/~nwdb/>) such that all non-words had orthographically existing onsets and bodies, as well as legal bigrams.

Design and procedure

The design and procedure was largely identical to that of Experiment 1 with the obvious difference that in this task participants were asked to indicate whether the targets (Japanese in the L2–L1 condition; English in the L2–L2 condition) were words or not. Item construction, including prime duration and SOA, was identical to that of Experiment 1. Two lists were

constructed in each priming condition in order to counterbalance across the prime factor. In the translation condition, targets that were preceded by their masked L2 translation equivalent on List A were preceded by a masked control prime on List B, and vice versa. In the within-L2 repetition priming condition, critical targets paired with a repetition prime on List A were paired with a control prime on List B, and vice versa. The repetition prime consisted of the same letter string as the target, but was presented in lowercase letters. The control prime was completely unrelated to the target except that it was matched with the target on length, frequency, and concreteness. No target word or prime word was repeated within lists. Items were presented randomly to each participant and RTs were recorded to the nearest millisecond.

Results

The same trimming procedure used in Experiment 1 was employed again. In the translation condition, mean response times were 529 ms for Japanese targets preceded by their English (L2) translation equivalents, and 525 ms for targets preceded by English control primes.

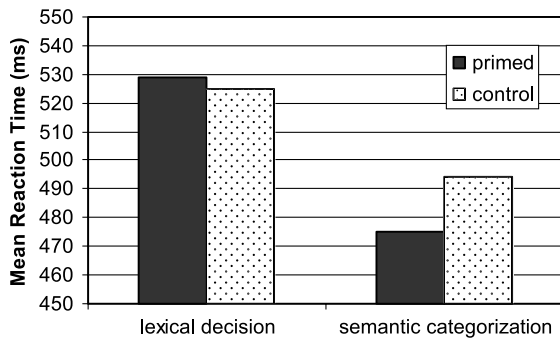


Fig. 1. Response latency means for L1 (Japanese) targets on Yes trials by task (lexical decision vs. semantic categorization) and priming condition [preceded by L2 (English) translation equivalent (primed) vs. preceded by an unrelated L2 word (control)].

Table 1

Mean lexical decision times (ms), percent error rates, and priming effects as a function of word type in within-L2 repetition priming (Experiment 2)

	Word Type			
	Abstract		Concrete	
	Low-frequency	High-frequency	Low-frequency	High-frequency
Control	990 (29)	659 (3)	860 (22)	630 (3)
Prime	913 (33)	590 (2)	737 (16)	567 (2)
Priming effect	77	69	123	63

This slight inhibitory effect of 4 ms was not significant (all F s < 1). The mean error rate was 5.6% on non-word trials and 1.3% on experimental trials. Error rates on experimental trials did not differ between conditions. Crucially for the purposes of this study, an ANOVA revealed a significant interaction between the priming effects of Experiments 1 and 2 ($F(1, 36) = 4.62$, $P = .03$; $F(1, 51) = 6.70$, $P = .012$). This interaction, depicted in Fig. 1, provides convincing support for a task difference, with reliable priming in semantic categorization but not lexical decision.

In the within-L2 repetition priming condition, two of the low-frequency abstract items had to be excluded from analysis because of exceptionally high-error rates. In order to balance the design, two items were randomly selected and removed from each of the other word types. The mean lexical decision times and percent error rates are presented in Table 1. For each comparison of interest, two analyses of variance were performed, one treating subjects as a random factor ($F1$), the other treating items as a random factor ($F2$). The factors were Word Type (abstract vs. concrete), Frequency (high vs. low), and Prime Type (identity vs. unrelated). The factor of Prime Type was a repeated measures factor in both analyses, but the factors of Frequency and Word Type were repeated in the subject analysis but not in the item analysis.

As Table 1 shows, there were clear main effects of word type, frequency, and prime type. Participants exhibited a clear concreteness effect, responding on average 89 ms faster to concrete words than to abstract words. This difference was found to be highly reliable, $F(1, 17) = 15.48$, $P < .01$; $F(1, 68) = 12.67$, $P < .01$. Likewise, participants exhibited a clear frequency effect. Mean response times were 263 ms faster for high-frequency items than they were for low-frequency items. Again, this difference was found to be highly reliable: $F(1, 17) = 88.70$, $P < .01$; $F(1, 68) = 174.19$, $P < .01$. Crucially, participants exhibited a clear repetition priming effect (83 ms), which was found to be reliable in both the subject and items analyses, $F(1, 17) = 18.95$, $P < .01$; $F(1, 68) = 25.67$, $P < .01$. None of the interactions between factors reached significance. These findings reveal robust

lexical processing skills and, hence, indicate that the lack of L2–L1 priming in lexical decision was not due to participants' inability to process L2 primes under these particular masking conditions.

Discussion

Taken together, the findings of Experiments 1 and 2 demonstrate clearly that L2–L1 priming is task dependent. Despite observing robust within-L2 masked repetition priming (confirming the findings of Gollan et al., 1997; Jiang, 1999), and clear L2–L1 masked translation priming in semantic categorization (confirming the finding reported by Grainger & Frenck-Mestre, 1998), there was a marked absence of L2–L1 translation priming in the lexical decision task (cf. Gollan et al., 1997; Grainger & Frenck-Mestre, 1998; Jiang, 1999; Keatly et al., 1994). Models of bilingual lexical processing such as the RHM have difficulty accounting for this task dependency because L2 form–meaning connection strengths, proposed to be the limiting factors in L2–L1 priming, cannot at the same time limit priming in one task but not another.

One may want to argue, though, that the task difference is due simply to the fact that consideration of meaning is critical for semantic categorization but not for lexical decision. That is, participants essentially ignore meaning when they make lexical decisions, but not when they make semantic category decisions. The problem with this argument is that it fails to explain the existence of strong semantic and associative priming effects in lexical decision (see Neely, 1991 for a review), which suggests that semantic and conceptual properties are highly relevant to the decision process. Alternatively, it could be argued that the semantic categorization task itself amplifies the strength of the form–meaning connections, so that the weak L2 form–meaning connections are now strong enough to produce priming. If this were the case, then we should observe priming for both exemplars (as was the case in Experiment 1) and non-exemplars. An alternative possibility is that in the semantic categorization task, the existence of the category itself might act as a context, which may serve to emphasize the category-relevant semantic properties of both the prime and target, thereby increasing the semantic overlap between them. If this were the case, then priming would be expected only for exemplars, since the non-exemplars would not have any category-relevant semantic features, and hence would be unaffected by the category context. Thus, it is critical to determine whether priming is obtained for non-exemplars as well as exemplars: a simple amplification model predicts priming for both, but a context selection model predicts priming only for exemplars. We test this possibility in Experiment 3.

Experiment 3—semantic categorization with non-exemplars

Experiment 3 was identical to that of Experiment 1, the only exception being that the critical items from Experiment 1 were now presented as non-exemplars. This was achieved by randomly selecting 48 of the original 52 items and distributing them evenly between 6 new categories (*military title, part of speech, weapon, natural earth formation, weather phenomenon, and girl's first name*). These particular categories were chosen for two reasons: (1) none of our critical items could be construed as having exemplar status in any of the categories and (2) none of the exemplars in these categories were cognates with their English translation equivalents. Each category consisted of 30 items, 15 exemplars and 15 non-exemplars. In order to maintain consistency with Experiment 1, the critical items ($N = 8$ per category) were interleaved with an equal number of exemplars and presented randomly within categories, with the only constraint being that the critical items (and an equal number of exemplars) were presented subsequent to the other items in each category. The participants ($N = 20$) were recruited from the same student population of individuals tested in Experiments 1 and 2. Again, each participant was paid for their participation.

Results

The same trimming procedure used in Experiments 1 and 2 was employed again. Mean response times for the critical items as non-exemplars were 605 ms when preceded by English (L2) translation-equivalent primes, and 609 ms when preceded by English control primes. This slight priming effect of 4 ms was not found to be significant (all $F_s < 1$). Error rates ($M = 3.5\%$) did not differ significantly between conditions. Importantly, there was an interaction between the priming effects of Experiment 1, when the critical items were exemplars, and those of the present experiment, when the critical items were non-exemplars ($F(1, 36) = 6.99, P = .01$; $F(1, 92) = 3.53, P = .07$).

Discussion

The lack of a reliable priming effect in the present experiment for non-exemplars suggests that the requirements of the semantic categorization task alone do not simply strengthen the form–meaning connections necessary to produce L2–L1 priming. Furthermore, the nature of the interaction between the priming effects observed in Experiment 1 and the lack of priming in the present experiment suggests that reliable priming in the L2–L1 direction is observed in semantic categorization only when participants generate decisions for exemplars. This result is consistent with the proposal that the

category acts as a kind of context which emphasizes the category-relevant properties of the prime and target. Such an effect would be expected only for exemplars of the category. In what follows, we provide a possible explanation for this pattern of findings.

The Sense Model

We begin with an issue that has not received adequate attention before in the cross-language priming literature: most words are polysemous and the range of senses that a word can have tends to be language specific.³ For example, the English word “black” and the Japanese word “*kuroi*” are translation equivalents, despite having little besides the one sense COLOR in common. According to WordNet 1.6 (Fellbaum, 1998), the word “black” has 21 different senses. For example, “black” is used to refer to a type of humor, as well as a calamitous day on Wall Street (*black Monday*). In Japanese, “*kuroi*” can be used to refer to those who are evil-minded (*black belly*), as well as those who are well tanned or guilty of a crime. Hence, the language-specific shades of meaning, or senses, for these two words extend well beyond the single sense determining translation equivalency. The relevance of this for translation priming is that the amount of priming may depend not only on the overlap in the semantic senses activated by the prime and the target, but, crucially, on the ratio of primed to unprimed senses associated with the target.

In thinking about how to formulate an account of the translation priming asymmetry that makes central the fact of polysemy, we have begun with a set of assumptions that, though fundamentally different from those on which the Revised Hierarchical Model and the Separate Memory Systems Model are based, resonate fairly closely with those of the Distributed Conceptual Feature Model (de Groot, 1992). In de Groot’s distributed conceptual feature model (DCFM), lexical nodes are associated with a distributed set of conceptual features, as opposed to localist conceptual representations. According to the DCFM, there are varying degrees of overlap of conceptual features between L1 and L2 meanings, depending on what type of word is represented. For example, de Groot (1992, 1993) has argued that there is more featural overlap between translation equivalents for concrete words than abstract words because concrete referents tend to have the same shape, size, and function cross-linguistically. The DCFM is

intuitively appealing on several fronts, not least of which is its straightforward account of how translation equivalents can have language specific meanings. In fact, de Groot (1993) has referred to the DCFM as a “mixed” model because it is possible for the meaning representations to range from having complete featural overlap (essentially the architecture of a compound bilingual) to having no featural overlap (i.e., the architecture of a coordinate bilingual) depending upon the type of word in question.

A critical problem with the DCFM, as currently specified, though, is its difficulty in accounting for asymmetries in translation performance. That is, because this model attributes translation performance to the conceptual features common to both L1 and L2, and because the amount of featural overlap between two translation equivalents is constant regardless of translation direction, any effect (such as priming) found in one direction should occur equally strongly in the other direction. As such, the DCFM has a particular difficulty in accounting for the masked translation priming asymmetry. In their revision of the conceptual feature model, Kroll and de Groot (1997) have addressed this shortcoming by proposing weak connections between L2 lexical nodes and conceptual features. As the findings of Experiments 1–3 indicate, though, the connection-strength approach may not be the correct one. Instead, we have chosen to pursue a different modification to the DCFM, one in which lexical semantic representations are assumed to be bounded and comprised of distinct “bundles” of features corresponding to distinct usages, which we refer to as semantic senses. These modifications to the DCFM present us with the hypothesis that what is critical in observing translation priming is the degree to which the complete lexical semantic representation (as opposed to just the features in common between translation equivalents) has been activated by the prime.

The number of bundled features, or senses, present in a lexical semantic representation can range from one to several dozen. The exact number depends on how many usages that particular word has and, important to the discussion of the bilingual lexicon, how knowledgeable the individual is of those usages. Because the shades of meaning and range of usages that a particular word may have vary from language to language, it is assumed that several of a word’s senses (i.e., the way in which conceptual features are bound together) will be language specific, with an obvious exception for the semantic sense(s) determining translation equivalency. The semantic sense determining translation equivalency is shared between L1 and L2 lexical entries, and is thus considered to be largely, if not completely, identical between L1 and L2 (see Fig. 2).

By assuming that each sense of a word constitutes a distinct mental representation within a lexical semantic

³ Here, we distinguish polysemy from ambiguity. A polysemous word (e.g., “call”) has several shades of meaning or usages (i.e., senses) that are all considered to be part of the same word (e.g., “telephone call” and “to call out”). An ambiguous word like “bark,” though, can be considered to be two distinct words (“to bark” and “tree bark”) that just happen to share the same orthographic and phonological form.

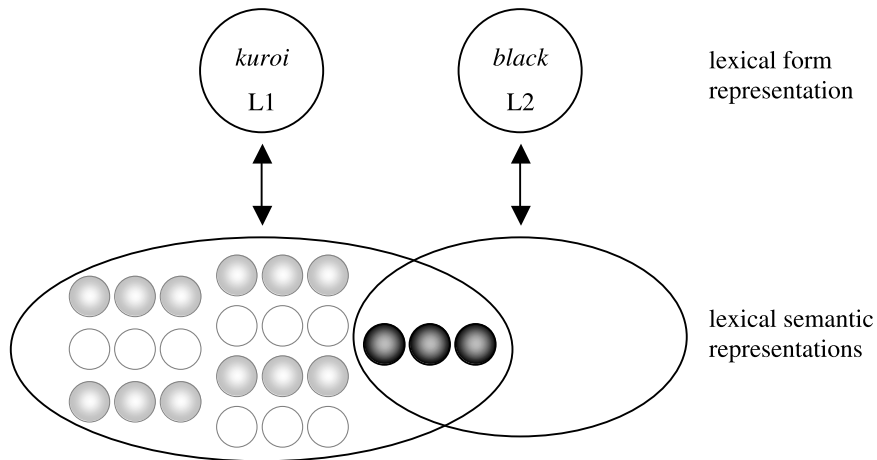


Fig. 2. A schematic representation of two translation equivalents according to the Sense Model. The semantic sense determining translation equivalency (and membership in the category COLOR) is depicted in dark grey and is shared between translation-equivalent lexical semantic representations. The language-specific senses are depicted in light grey or white. By hypothesis, the proportion of activation realized throughout the lexical semantic representation is relevant in lexical decision, which leads to a priming asymmetry whenever a representational asymmetry is present (i.e., the proportion of activation in this example is 1/8 in the L2–L1 direction, whereas it is 1/1 in the L1–L2 direction). When decision making is constrained semantically, as it is thought to be in semantic categorization, only activation of the sense determining category membership is relevant, which leads to largely symmetrical priming (i.e., the proportion of activation realized across the relevant sense is 1.0 in both L1–L2 and L2–L1 directions).

representation, we are presented with the possibility of a representational asymmetry between related words. Instances of this asymmetry can be found between words within a language as well as between translation equivalents across languages. Taking a word pair within English as an example, a representational asymmetry is said to exist between two lexical representations when (one) they share at least one semantic sense and (two) when one word has many different senses (e.g., “head” has multiple usages, including: *part of an animal’s body*; *head of state*; *full head of steam*; *heading in a direction*, etc.) and the other has very few (e.g., “skull” only has one sense: *part of an animal’s body*). Within languages, it is relatively uncommon to find word pairs that meet these criteria. We suggest that translation equivalents, on the other hand, frequently meet the criteria of a representational asymmetry. For example, translation equivalents, by virtue of being “equivalent,” presumably share at least one semantic sense. Furthermore, because bilinguals, generally speaking, are more proficient in their L1 than in their L2 and, thus, are more familiar with the range of usages that L1 words may have, the number of senses associated with L1 words is thought to exceed the number of semantic senses associated with L2 words.

The translation priming asymmetry in lexical decision is seen to be the natural consequence of this representational asymmetry. Recent work investigating the influence of semantic senses on lexical decision processes has found that response times are sensitive to the

number of senses associated with a particular target (Rodd, Gaskell, & Marslen-Wilson, 2002; see also General discussion). Given this finding, it is reasonable to consider the possibility that semantic priming reflects the efficacy of the prime word to preactivate the semantic senses associated with the target word (cf. Cree, McRae, & McNorgan, 1999, for a similar “feature overlap” account of semantic priming). According to this account of semantic priming, the amount of priming would increase as the proportion of primed to unprimed senses in the lexical semantic representation of the target word increases. In the L1–L2 direction, we propose that the proportion of L2 senses primed by the L1 prime is going to be very high, if not complete. This is because most, if not all, of the senses associated with the L2 form are also associated with the L1 equivalent form (much like the sense of “skull” is also associated with “head”). Priming in this direction, then, should be very reliable, and it is (de Groot & Nas, 1991; Gollan et al., 1997; Jiang, 1999; Keatly et al., 1994; Williams, 1994). However, in the reverse direction (L2–L1), the proportion of L1 senses primed by the L2 prime is going to be very low (e.g., 1:8). This is because it is frequently the case that there are many senses associated with the L1 form that are not similarly associated with the L2 prime (again, much like the relationship between “head” and “skull,” where many senses of “head” are not associated with “skull”). Accordingly, priming in this direction should be very weak, and it is (Gollan et al., 1997; Grainger & Frenck-Mestre, 1998;

Jiang, 1999; Jiang & Forster, 2001; Keatly et al., 1994; Sanchez-Casas et al., 1992; as well as Experiment 2 of the present study).

If it is correct to assume that L2–L1 priming does not occur (or is very weak) in lexical decision because an insufficient number of L1 senses are preactivated by the L2 prime, then there must be something about the semantic categorization task that serves to restrict the number of senses recruited when a decision is generated. As we have already indicated, it is plausible to argue that the category acts as a context that serves to focus how decisions are made in response to the target, so that only the semantic sense that is relevant to the category is taken into account (e.g., in order to decide that “black” is a color, one does not need to take into consideration the multiple senses that “black” may have). Essentially, we argue that the semantic categorization task turns many-sense targets into one-sense targets by “filtering” out category-irrelevant senses from the decision making process. Assuming that the prime–target pairs have been properly selected, this suggests that the L2 prime will be more effective in priming the L1 target because the only sense that is relevant for the decision will be the very sense that has been primed by the L2 prime.

The Sense Model differs from other models of bilingual lexical representation and processing in that its explanation of the translation priming asymmetry is not restricted to a peculiarity of the bilingual lexicon (e.g., weak connection strengths). That is, the Sense Model predicts the same pattern of priming effects (and task differences) for word pairs within a language whose relationship meets the representational asymmetry criteria specified above. In fact, there are reasons to think that the strongest test of the Sense Model would be to use within-language word pairs in a masked priming experiment with lexical decision and semantic categorization. If the assumptions of the Sense Model are correct, then word pairs like “head–skull,” which effectively instantiate the assumed representational asymmetry between translation equivalents, should produce the same priming asymmetry in lexical decision but not semantic categorization. These predictions of the Sense Model are tested in the following experiments.

Experiment 4—within-language priming asymmetry in lexical decision

In the present experiment, we test the predictions of the Sense Model in a masked priming task with lexical decision. If the assumptions of the Sense Model are correct, English materials that effectively recreate the representational asymmetry thought to exist between translation equivalents should produce priming

in the many-to-one direction (e.g., “head–SKULL”), but not in the one-to-many direction (“skull–HEAD”).

Participants

Forty-four undergraduates at the University of Arizona participated for course credit. Twenty-two of the participants were tested in the many-to-one (thought to be analogous to L1–L2) direction (Experiment 4a) and 22 were tested in the one-to-many (thought to be analogous to L2–L1) direction (Experiment 4b). All participants were native speakers of English.

Materials

Several different selection criteria were employed. First, it was imperative to select word pairs that were thought to share a common sense. This was done by selecting synonyms, or close synonyms. Second, it was important that the materials satisfied the numerical discrepancy criterion of the representational asymmetry between the words. This was done by ensuring that one word in the pair had only one sense (e.g., “oar” has only one sense) and that the other word in the pair had several (e.g., “paddle” has 10 senses). Sense counts were taken from WordNet 1.6 (Fellbaum, 1998). One-sense words were paired with words having eleven senses on average (see Appendix A for a complete list of word pairs and sense counts). Forty word pairs in total were created. An equal number of non-words was generated by the ARC Nonword Database (Rastle, Harrington, & Coltheart, 2002) and interleaved with the experimental items. Additionally, 40 words that were matched in length and frequency with the experimental primes were selected to serve as control primes.

Design and procedure

Experiments 4a and 4b were identical in design and procedure. Participants in both experiments were asked to decide whether the target item was an English word or not. In Experiment 4a, the primes had many senses, while the targets had just one. This was reversed in Experiment 4b. Each trial consisted of the following sequence of events, each one following immediately after the other: forward mask (#####) for 500 ms, prime in lowercase letters for 41 ms, and target in uppercase letters for 500 ms. Two lists were constructed for both Experiments 4a and 4b in order to counterbalance across the prime factor. Targets that were preceded by their related prime on List A were preceded by a control prime on List B and vice versa. Items were presented randomly to each participant and RTs were recorded to the nearest millisecond.

Results

The same trimming procedure used in Experiments 1–3 was employed here. In Experiment 4a, the many-to-one (or “L1–L2”) condition, mean response times were 550 ms for targets (e.g., *OAR*) preceded by a related prime (e.g., *paddle*) and 573 ms for targets preceded by a control prime. An ANOVA showed that this facilitation effect of 23 ms was reliable ($F(1, 20) = 12.41, P = .002$; $F(1, 38) = 14.26, P < .001$). The mean error rate was 5.9% and was not different between experimental and control conditions (all $F_s < 1$).

In Experiment 4b, the one-to-many condition (or “L2–L1”) direction), the pattern of results was markedly different despite using the same word pairs, albeit in the opposite direction. Mean response times were 538 ms for targets (e.g., *PADDLE*) preceded by a related prime (e.g., *oar*) and 528 ms for targets preceded by a control prime. An ANOVA showed that this inhibitory effect of 10 ms was not reliable ($F(1, 20) = 1.71, P = .21$; $F(1, 38) = 3.24, P < .079$). The mean error rate was 5.7% and was not different between conditions (all $F_s < 1$). Importantly, there was a significant interaction between the results of Experiments 4a and 4b ($F(1, 40) = 10.49, P = .002$; $F(1, 72) = 15.70, P < .001$).

One possible explanation for this interaction that we consider has to do with the large difference in surface frequency between the primes and the targets. Many-sense words, by virtue of having several different usages, have a much higher surface frequency than few-sense words. Because surface frequency and number of senses are so highly correlated, we were unable to control appropriately for this in our item selection, and the concern is that high-frequency words may be better semantic primes than low-frequency words. Although it is well known that frequency rarely interacts with masked repetition priming effects (cf. Experiment 2), much less is known about the possibility of frequency interacting with masked semantic priming effects. In order to address this possibility, we calculated the priming effect for each item from Experiments 4a and 4b and entered them into a simultaneous multiple regression analysis with the sense counts of the target words and the surface frequencies of the prime words included as predictors. This analysis revealed that although the sense counts of the target words accounted for a significant proportion of the priming effect ($t = -2.97, P = .003$), the surface frequencies of the prime words did not ($t = -0.27, P = .788$).

Discussion

These findings confirm the predictions of the Sense Model and hence provide support for the assumption that a representational asymmetry between L1 and L2

equivalents may be the source of the translation priming asymmetry. Masked priming was observed in the lexical decision task between many-sense words (e.g., “paddle”) and one-sense words (e.g., “oar”) in the many-to-one direction. This effect is analogous to the masked translation priming effects found with bilinguals in the L1–L2 direction (de Groot & Nas, 1991; Gollan et al., 1997; Jiang, 1999). Using the same word pairs, but now in the one-to-many direction, no priming was observed. This finding is analogous to the lack of masked translation priming in the L2–L1 direction (Gollan et al., 1997; Grainger & Frenck-Mestre, 1998; Jiang, 1999; Sanchez-Casas et al., 1992; and Experiment 2 above).

These results are important in at least two different ways. First, they provide just the kind of evidence needed to confirm the central assumption of the Sense Model. Second, they represent a relatively unusual case of semantic priming in that this is a *masked* semantic priming effect in lexical decision; most previous studies reporting semantic priming effects have used much longer prime durations or semantic judgment tasks.⁴

Experiment 5—symmetrical priming in semantic categorization

In Experiment 4, the central assumption of the Sense Model was confirmed, namely that the priming asymmetry frequently observed in translation priming is the consequence of a representational asymmetry between L1 and L2 translation equivalents. We now turn our investigation to the task difference in L2–L1 translation priming. Here, the question is how the semantic categorization task eliminates the priming asymmetry that is clearly present in lexical decision. We have suggested that in lexical decision, there is no task relevant category to restrict how participants generate decisions, effectively diluting the efficacy of the L2 prime. In semantic categorization, though, the presence of relevant category information may serve to focus how decisions are made in response to the target. The consequence of this “focusing” effect is that only the semantic sense that is relevant to the category is taken into account, thereby allowing the L2 prime to be effective. That is, by hypothesis, semantic categorization effectively turns many-sense words into one-sense words for the purposes of the task. If this hypothesis is correct, then priming should be symmetrical between many- and few-sense words in

⁴ Subsequent replications of this experiment have uncovered a curious fact: many-to-one priming effects are only observed when at least 50% of the positive items are of the “paddle–OAR” type. Currently, we have no explanation for this result, other than to point out that this may be similar to the relatedness proportion effects reported by Bodner and Masson (2001) in masked priming.

semantic categorization, despite being clearly asymmetrical in lexical decision (Experiment 4). We test this hypothesis in the present experiment.

The masking procedure adopted in this experiment differs from the previous cross-language semantic categorization experiments (Experiments 1 and 3) in that the duration of the interpolated mask is reduced from 150 to 13.7 ms. This was necessary because the primes were now presented in L1 rather than L2, and with a 150 ms post-prime mask, the prime may have been visible (i.e., prime visibility depends on language competence). The reason for including an interpolated mask at all was that a review of the literature revealed that this is the typical presentation sequence employed in masked semantic priming experiments using semantic categorization with native speakers (Bueno & Frenck-Mestre, 2002; Frenck-Mestre & Bueno, 1999). It also makes these within-language experiments more comparable to the earlier between-language experiments. However, because of concerns about the visibility of the primes presented in L1, we also included a test of prime awareness that was given after the experiment proper.

Participants

Forty undergraduate participants at the University of Arizona participated for course credit. Twenty of the participants were tested in the many-to-one (“L1–L2”) direction (Experiment 5a) and 20 were tested in the one-to-many (“L2–L1”) direction (Experiment 5b).

Materials

Items were selected in largely the same way that items for Experiment 1 had been selected. Forty-two words taken from 8 different semantic categories were selected to serve as many-sense (or “L1”) targets in the one-to-many priming direction. These were chosen according to two different criteria: one, they had to be among the first 10 exemplars of the category (Battig & Montague, 1969; Uyeda & Mandler, 1980) and, two, they had to have at least 5 semantic senses associated with them (WordNet 1.6). The semantic categories were as follows: *article of clothing*, *article of furniture*, *bird*, *carpenter’s tool*, *color*, *four-footed animal*, *kitchen utensil*, *body part*. Additionally, 42 words were selected to serve as one-sense (“L2”) primes. Selection of these prime words was done according to the same criteria used in Experiment 4: that is, items had only one sense (WordNet) and that sense had to be shared with the target. An example of a prime–target pair is “skull–HEAD.” Here, “skull” only has one sense, which is shared with “head.” “Head,” on the other hand, has more than 20 senses, only one of which constitutes the meaning of “skull.” In addition to these 42 prime–target pairs, additional exemplars were selected for a total of 10 exemplars per category. These

filler items were also from the first 10 exemplars for each category, but the number of senses associated with them was not controlled for. Eighty non-exemplars were also selected. These were matched with the exemplars on length and frequency, but care was taken to ensure that none of these items could be construed as having exemplar status in any of the categories in the experiment. Finally, 160 additional words were selected to serve as prime words for the filler exemplars, the non-exemplars, and for the critical targets in the control condition. These were matched with the experimental primes on length and frequency, but again could not be construed as having exemplar status in any of the categories present in the experiment.

Design and procedure

The design and procedure was identical for Experiments 5a and 5b. Much like Experiment 1, items were blocked according to semantic category. For each category, there were 10 exemplar and 10 non-exemplar trials. Items were presented in a different random order for each participant with the constraint that the critical items followed the filler items in each category. Critical targets (e.g., “HEAD”) were matched with a related prime (e.g., “skull”) as well as a control prime (e.g., “shop”). Two lists were constructed for both Experiments 5a and 5b in order to counterbalance across the prime factor. Targets that were preceded by their related prime on List A were preceded by a control prime on List B and vice versa. As in Experiments 1 and 3, each trial consisted of the following sequence of events: first, a forward mask (#####) was presented for 500 ms, followed by the prime word (related or control) in lowercase letters for 41 ms, followed by a backward mask for one refresh cycle (13.7 ms), and then the target word in uppercase letters for 500 ms. Because of concerns about the visibility of the primes presented in L1, we also included a test of prime awareness that was given after the experiment proper. This test required participants to indicate whether or not the prime word contained the letter “e.” Participants were first shown several practice trials at a much slower presentation rate so that they could see the prime word. After it was clear that they understood the requirements of this new task, they were presented with 42 trials in a different random order for each participant. Half of these trials had prime words that contained the letter “e,” and half did not. The targets were completely unrelated words, and none of the targets contained the letter “e.” Thirty-two of the trials were presented using the same temporal characteristics used in the experiment proper; on the 10 remaining trials (5 with the letter “e” and 5 without), the prime word was presented for 10 refresh cycles (137.3 ms). On these trials, the prime word was visible, though only very briefly. The purpose of presenting

trials with visible primes was to encourage participants to continue trying to do the task. An acceptable accuracy rate on these trials allows us to safely assume that participants were performing the task to the best of their abilities.

Results and discussion

Experiment 5a—*one-to-many in semantic categorization*

The same trimming procedures used in the previous experiments were again employed here. In Experiment 5a, mean response times in the one-to-many (or “L2–L1”) direction were 473 ms for targets (e.g., *HEAD*) preceded by related one-sense primes (e.g., *skull*) and 486 ms for the same targets when preceded by a control prime (e.g., *shop*). An ANOVA showed that this facilitation effect of 13 ms was reliable ($F(1, 18) = 8.15$, $P = .01$; $F(1, 40) = 7.02$, $P = .01$). Furthermore, just as there was an interaction between the results of the semantic categorization and lexical decision tasks in translation priming (Experiments 1 and 2), so was there an interaction between the results of the present experiment and those of Experiment 4b ($F(1, 36) = 6.54$, $P = .01$; $F(1, 76) = 5.47$, $P = .02$). These results confirm the task difference between semantic categorization and lexical decision in one-to-many priming. In so doing, these findings provide further support for the hypothesis that semantic categorization serves to restrict the amount of semantic information recruited to generate a decision, thereby enhancing the effectiveness of the few-sense prime word.

The mean error rate in this experiment was 5.4%: 3.3% on the trials with related primes and 7.5% on trials with control primes. This difference was significant in both the subject and items analyses ($F(1, 18) = 7.06$, $P = .01$; $F(1, 40) = 7.25$, $P = .01$), indicating that participants benefited from a related prime in terms of both reaction times and accuracy. The error rates constitute the only difference between this experiment and Experiment 1. In Experiment 1, bilinguals did not exhibit any difference in error rates between conditions. Here, on the other hand, there is the suggestion that information presented in the prime stimulus was not only activating the target representation, but also entering into the decision making process. Although it is somewhat unclear how this could occur (participants were unaware of the prime, see below), the pattern of errors speaks to the possibility of a decision conflict. That is, participants may have begun generating a decision on the basis of the prime’s category membership, which would facilitate decisions on the experimental trials, but not on the control trials. This is a possibility that we consider further in Experiment 6.

With respect to participants’ performance on the prime awareness task, the masking procedure appears to have been effective in preventing awareness of the prime.

On trials in which the presentation sequence was the same as in the semantic categorization task, participants correctly detected prime words with the letter “e” 51.5% of the time. Their performance did not differ from chance ($F < 1$). On trials in which the prime was presented for 10 refresh cycles, participants correctly identified prime words with the letter “e” 87.8% of the time. Taken together, these findings indicate: (1) that participants were performing the “e-detection” task to the best of their abilities and (2) that on the trials in which prime duration matched the prime duration used in the semantic categorization task (41 ms), participants were generally not able to detect the prime. As such, these findings rule out awareness as a possible source of the priming effect observed in Experiment 5a.

Experiment 5b—*many-to-one in semantic categorization*

Although Experiment 5a confirms the task difference in the one-to-many direction, and, hence, provides strong support for the assumptions of the Sense Model, our account of the task difference predicts symmetrical priming in semantic categorization. For the sake of completeness, then, we conducted Experiment 5b to confirm that many-to-one priming occurs both in lexical decision and semantic categorization. In this experiment, mean response times were 532 ms for targets (e.g., *SKULL*) preceded by related many-sense primes (e.g., *head*) and 552 ms for the same targets when preceded by a control prime. This 20 ms effect was found to be significant in both the subject and items analyses ($F(1, 18) = 7.28$, $P = .01$; $F(1, 40) = 7.67$, $P = .008$). Importantly, this effect does not differ from the 23 ms priming effect found in the many-to-one direction in lexical decision (Experiment 4a). Again, these results confirm the predictions of the Sense Model. According to the Sense Model, there should be no task difference when targets have very few senses. This is because the proportion of the target’s senses activated by the prime will be very high, which will facilitate decisions in both lexical decision and semantic categorization. In other words, the “focusing effect” of semantic categorization does not provide any benefit when the target has only one sense to begin with.

The mean error rate in Experiment 5b was 8.5%. Again, there was a difference between experimental and control trials. On trials in which the target was preceded by a related prime, the error rate was 6.02%; the error rate on the control trials was 10.93%. This difference was very close to reaching significance in the subject analysis and did reach significance in the items analysis ($F(1, 18) = 3.82$, $P = .07$; $F(1, 40) = 5.04$, $P = .03$). Again, this difference suggests that participants were able to benefit from the prime both in terms of reaction times and accuracy. One possible interpretation is, as suggested above, participants were initiating a response based on the category membership of the prime. If this

was the case, then the facilitation effect would be a congruency effect, not a priming effect. Again, this is a possibility that we address in Experiment 6.

Turning now to the prime awareness task, the participants in Experiment 5b did not do any better than those in Experiment 5a. The mean accuracy rate on trials with a 41 ms prime duration (as in the semantic categorization task) was 53.2%. This did not differ significantly from chance ($F = 2.5$; $p = .11$). Participants did much better on trials where the prime was presented for 10 refresh cycles (137.3 ms), correctly identifying the prime words containing the letter “e” 83% of the time. Again, this pattern of findings indicates very clearly that: (1) participants were trying to do this task and (2) that despite their efforts, they were not able to perceive the prime when the presentation sequence was identical in its temporal characteristics to that used in the semantic categorization task. Following from this, it is unlikely that the priming exhibited by these same participants in the semantic categorization task could have been due to their awareness of the prime’s presence.

Experiment 5 has contributed in an important way to our understanding of the task difference between semantic categorization and lexical decision in translation priming. It is clear from this experiment that semantic priming occurs equally in both the one-to-many and the many-to-one direction, despite the representational asymmetry between prime and target. This is quite different from the results seen in lexical decision. In lexical decision, priming only occurs in the many-to-one direction. Although this pattern of findings fits very well with the predictions of the Sense Model, there is a possible alternative explanation for the task difference, which we address now. This is the possibility of a decision conflict.

This possibility was pointed out by Davis, Kim, and Sanchez-Casas (2003) in their critique of the original finding reported by Grainger and Frenck-Mestre (1998). They argued that the control condition used in the Grainger and Frenck-Mestre (1998) experiment (the same as in our Experiments 1 and 5) may have been inappropriate. For the critical items, the unrelated prime was a non-exemplar, while the target was an exemplar. This may have generated a decision conflict for unrelated items, but not for related items (since both prime and target were exemplars), which would produce a congruence effect in the place of a true priming effect. Previous attempts to discover such congruence effects, however, have not been successful. In a recent experiment, Damian (2001) found that decision conflicts in a size judgment task arose only when primes had previously appeared as targets. Damian interpreted those findings to mean that decision conflicts, when they arise, were the result of automatized stimulus–response mappings, not unconscious categorization of the prime. Because none of the primes in the present set of

experiments ever appeared as targets, there is no reason to believe that our participants ever developed such automatized stimulus–response mappings. In a separate investigation of the same issue, Forster et al. (2003), using the category ANIMAL, found no difference between congruent pairs (e.g., “shark–ROBIN”) and incongruent pairs (e.g., “badge–ROBIN”), suggesting that congruence effects could safely be ignored. However, this was true only for exemplars. For non-exemplars, a congruence effect (15 ms) was detected. Similar results for exemplars were obtained by Bueno and Frenck-Mestre (2002), who reported similar priming effects for “dolphin–WHALE” when the control condition used non-exemplars as primes (e.g., “helmet–WHALE”) and when the control condition used congruent, but semantically dissimilar, exemplars as primes (e.g., “sparrow–WHALE”). Since neither of these experiments report evidence of a congruence effect for exemplars, a decision conflict account for the priming effects seen in Experiments 1 and 5 seems very unlikely. However, recent evidence (Forster, in press) indicates that very substantial congruence effects can be generated for both exemplars and non-exemplars when a small category is used (e.g., *farm animal*), but not when a large category is used (e.g., *animal*). Since some of the categories used in Experiments 1 and 5 might be described as “small,” it seemed wise to eliminate this possible explanation of the priming effect. Accordingly, we address this possibility in Experiment 6 by conducting a replication of Experiment 5. The only difference between the two experiments was that in Experiment 6 both experimental and control primes shared exemplar status with the target.

Experiment 6—semantic categorization with exemplars as control primes

Participants

Twenty-four undergraduate students from the University of Arizona participated for course credit.

Materials

The materials were identical to those used in Experiment 5 except for the control primes on the 42 critical trials. These were replaced so that both prime and target were exemplars of the same category. In selecting the new control items, care was taken to ensure that the control primes were semantically distant from the targets. For example, in the *BODY PART* category, the target “HAND” was paired with the experimental prime “wrist” and the control prime “kidney.” If participants are generating a response based on the category membership of the prime word, then there should be no difference in response times between experimental and

control trials. If, though, the prime stimulus is *priming* the target by activating a semantic sense in common between the prime and target, participants should be significantly faster on the experimental trials.

Design and procedure

The experimental design and procedure was identical to that of Experiment 5, except that we only tested in one direction. Because the source of the task difference between lexical decision and semantic categorization in the one-to-many direction is the critical question that this experiment addresses, it was only necessary to test in this direction. A significant priming effect in the one-to-many direction when using exemplars as control primes would allow us to rule out the possibility of a decision conflict.

Results and discussion

The same trimming procedures used in the previous experiments were used again here. Contrary to the predictions of the decision conflict account, semantic priming in the one-to-many direction was not affected by the presence of congruent, but semantically dissimilar, exemplars as primes. Mean response time for targets (e.g., “HAND”) preceded by experimental primes (e.g., “wrist”) was 463 ms; and mean response time for the same targets preceded by control primes (e.g., “kidney”) was 475 ms. This priming effect was found to be reliable in both the subject and items analyses ($F(1, 22) = 5.79, P = .02$; $F(1, 40) = 6.25, P = .01$). These findings confirm those that have been reported on at least two occasions now (Bueno & Frenck-Mestre, 2002; Forster et al., 2003). That is, masked priming effects for exemplars in semantic categorization are similar in magnitude regardless of whether the control condition includes category congruent items or not. As such, we can conclude that the task difference between lexical decision and semantic categorization in the one-to-many direction (and by extension, in the L2–L1 direction) is not attributable to a decision conflict in semantic categorization. Rather, these findings suggest that the task difference is due to the fact that the semantic categorization task serves to focus how decisions are made, so that only the semantic sense that is relevant to the category is taken into account. Taking “hand” as an example, one does not need to take into consideration the multiple senses that “hand” may have in order to decide whether it is a body part or not. Essentially, we argue that the semantic categorization task turns many-sense targets into one-sense targets by “filtering” out category-irrelevant senses from the decision making process, and that this enhances the efficacy of the one-sense prime. This issue will be considered again in the General discussion.

General discussion

The research here establishes several findings that, when taken together, present an important challenge to models of bilingual lexical representation and processing. These findings are: (1) asymmetrical translation priming in lexical decision, (2) symmetrical translation priming in semantic categorization, (3) robust within-L2 masked priming effects, and (4) the same pattern of findings for within-language word pairs that have the same representational asymmetry assumed between translation equivalents (see Table 2). As we have seen, models of bilingual lexical processing that propose limiting factors such as weak L2 form–meaning connection strengths (e.g., The Revised Hierarchical Model), or that the L2 lexicon may be stored in a separate memory module from the L1 lexicon (as in the separate memory systems account), cannot account for these findings. In what follows, we discuss in more detail how the Sense Model provides a parsimonious explanation for each of these findings in turn.

Masked translation priming asymmetry in lexical decision

The account of the priming asymmetry given by the Sense Model follows directly from the assumptions that it makes with respect to how lexical semantic information is represented. As we mentioned in the discussion of Experiment 3, the Sense Model, much like the distributed conceptual feature model (DCFM) (de Groot, 1992; de Groot, Dannenburg, & van Hell, 1994; Kroll & de Groot, 1997), holds that lexical (form-level) representations map onto distributed lexical semantic representations, as opposed to localist conceptual representations. Unlike the DCFM, though, a central assumption of the Sense Model is that semantic features are bundled into semantic senses within distinct lexical semantic representations. Following from this, the Sense Model assumes that masked translation (and semantic) priming is attributed not only to the overlapping semantic features between prime and target, but also to the ratio of primed to unprimed senses. When the semantic features in common between prime and target constitute a large proportion of the target’s lexical semantic representation, priming should be observed. On the other hand, when the semantic features in common between prime and target constitute only a small proportion of the target’s lexical semantic representation, the amount of priming may be too small to affect decision making in any significant way.

Why should we assume that lexical decision is sensitive to the proportion of semantic senses activated for a given target? Some support for this assumption comes from recent work by Rodd et al. (2002). Rodd et al. (2002) report that the controversial claim that there is an ambiguity advantage (i.e., that words with many

Table 2

Comparison of priming effects and error rates (in parenthesis) for all-English prime–target pairs by direction (many-to-one vs. one-to-many) and by task (lexical decision vs. semantic categorization) in Experiments 4–6

		Prime condition		Priming effect
		Prime	Control	
Lexical decision task	Many-to-one (Experiment 4a)	550 (5.70)	573 (5.75)	23
	One-to-many (Experiment 4b)	538 (5.1)	528 (6.2)	–10
Semantic categorization task	One-to-many (Experiment 5a)	473 (3.3)	486 (7.5)	13
	Many-to-one (Experiment 5b)	532 (6.0)	552 (10.9)	20
Semantic categorization task with exemplars as control primes	One-to-many (Experiment 6)	463 (3.3)	475 (5.5)	12

meanings are responded to faster than unambiguous words) is actually a *sense* advantage. They found that lexical decision times were actually slower for ambiguous words (words with distinct meanings, like “bank”) than for unambiguous words (though not significantly slower), whereas decision times for words with many senses (e.g., “hammer”) were significantly faster than for words with few senses (e.g., “cake”). All words were matched for frequency, length, familiarity and concreteness. Similar findings were obtained by Finkbeiner (2002) using a larger and randomly generated set of items.⁵ This involved using WordNet (Fellbaum, 1998) to generate lists of few-sense words (no more than two senses) and many-sense words (15 or more senses). Eighty items appearing in a frequency band of 20–80 per million were then randomly selected from each list of possible words for a total of 160 experimental items (items had a mean frequency of 45.8 and a mean length of 5.1 letters on each list). Using these items in a lexical decision task, a reliable difference between response times for many-sense words (533 ms) and few-sense words (564 ms) was found ($F(1, 14) = 14.44, P = .001$; $F(1, 78) = 36.54, P < .001$). Taken together, these findings suggest that lexical decision times are sensitive to the number of semantic senses associated with a particular target. It logically follows from this that in order to show priming, a large proportion of a target word’s semantic senses must be preactivated by the prime.

In summary, the masked translation priming asymmetry is argued to be the straightforward consequence of a representational asymmetry between L1 and L2 lexico-semantic representations. L1 words are argued to have relatively more senses associated with them than L2 words, and translation-equivalent words presumably only share a restricted number of senses. Following from this, it is argued that the proportion of L1 senses primed by an L2 equivalent is going to be lower than the pro-

portion of L2 senses primed by an L1 equivalent, which leads to the often observed translation priming asymmetry in lexical decision.

Task differences in masked translation priming

We return now to the question of how semantic categorization might eliminate the priming asymmetry. If it is correct to assume that L2–L1 priming does not occur (or is very weak) because an insufficient number of L1 senses are preactivated by the L2 prime, then there must be something about the semantic categorization task that serves to restrict the number of senses recruited when a decision is generated. Essentially, we argue that the semantic categorization task turns many-sense targets into one-sense targets by “filtering” out category-irrelevant senses from the decision making process. That is, in semantic categorization, the sense that triggers a decision will always be the sense that is relevant to the category, and this allows the L2 prime to be more effective. Assuming that the materials have been designed appropriately, the category-relevant sense will be the sense that both the L2 prime and the L1 target share, meaning that the sense that triggers a decision will already have been activated by the L2 prime.

This account raises several interesting questions. First, how is the category able to select out the relevant sense? It is interesting to note that a similar phenomenon apparently occurs in the context of neighborhood effects (Forster & Hector, 2002). These investigators found that in a semantic categorization task (e.g., “Is it an animal?”), non-words that were one-letter different from many words (e.g., *walley*) were classified just as rapidly as non-words with few neighbors (e.g., *braln*), unless one of those neighbors happened to be an animal name (e.g., *turple*), in which case a substantial inhibitory effect was observed. Evidently, a spelling check was triggered for *turple* because of its similarity to a word (which happened to be an exemplar), but not for a non-word such as *cishop*, despite its similarity to a word. Although this is hardly a surprising result, it does raise the question of how the decision system knows whether to ignore the presence of a word neighbor. Indeed, non-words that had

⁵ Forster (2000) has argued that in this kind of research, it is imperative that items are selected randomly to avoid inadvertent experimenter bias.

a large number of neighbors, one of which was an animal (e.g., *goan*) took no longer than non-words which also had an animal neighbor, but very few other neighbors (e.g., *caedel*). The puzzle here is to explain how the existence of an animal neighbor could be detected without applying some kind of semantic test to each neighbor. Somehow the category-irrelevant neighbors were filtered out, leaving just the category-relevant neighbor. We propose that a similar mechanism operates in sense-selection; category-irrelevant senses are automatically filtered out, leaving just the category-relevant sense.

Although we prefer the explanation provided by the filter hypothesis, there are other interpretations that need to be considered. According to the Sense Model, the locus of the priming effect is at the level of decision making. When the semantic sense of the target responsible for triggering a semantic judgment has already been activated by the prime, less time is needed to generate a decision, leading to a savings (or priming) effect. According to this account, then, the critical stage in the process occurs well after the form-level properties of the target have been successfully retrieved. But this is clearly not the only point in the overall process at which a savings could occur. For example, priming may be attributed to facilitation in the initial stage of word form recognition. Alternatively, it may be that the critical stage in the priming process involves a degree of similarity in the categorization processes of both the prime and target. Below we consider these alternatives in more detail.

The first alternative explanation suggests that what is primed in the semantic categorization task is not a semantic sense, but rather, the form-level properties of the target. According to this account, priming occurs when the processes involved in interpreting the prime alter the state of the lexical entry for the target, so that the target's word form is recognized more rapidly. In order for this to occur, a top-down process is necessary, where activation flows from the semantic level back down to the form level (e.g., Besner & Smith, 1992). This account essentially assumes that priming is a *perceptual* effect, i.e., that a semantically related prime facilitates recognition of the target's form. However, if this were the case, then it would be difficult to explain why there is no L2–L1 priming in lexical decision, but there is in semantic categorization. Faster perception of the target should lead to faster responses, whatever the task.

Another possible explanation is that what gets primed is the categorization process itself. For example, it might be that a categorization process is initiated for the prime as well as the target. When the target is semantically very similar to the prime, it may mean that less effort is involved in reaching a decision concerning the target. So, in the case of an item such as “whale–DOLPHIN,” it could be that having just computed the meaning of “whale,” and classifying it as an animal, it is much easier to decide that “dolphins” must also be an animal, because the processes

involved in classifying these words are very similar. This analysis is different from the suggestion that priming is the simple consequence of a sequence of two YES–YES responses (i.e., a congruence effect), one to the prime and one to the target, and, hence, could account for the findings presented in Experiment 6, where we found robust priming for “wrist–HAND” (the category being *Body Part*) even when the control prime would also have led to a YES response (“kidney–HAND”). According to this analysis, priming is observed in the “wrist–HAND” (or L2–L1 translation) condition because near identity of reference of the words permits faster semantic categorization. This would not be true for the words “kidney–HAND” in the control condition. In essence, when both the prime and target share many of the same conceptual features, classifying the target is going to be faster. The appropriate analogy here is with a syllogism of the form “If X is a member of a category, and Y is very similar to X, then it is likely that Y will also be a member.” Arguably, the similarity of reference is not relevant in lexical decision because what is critical in that task is whether the stimulus is a correctly spelled word. This, then, captures an important aspect of the task difference that we observed between semantic categorization and lexical decision. But there is also good reason to question this analysis. If the similarity in the categorization process between prime and target is the source of a time savings for exemplars, it should also lead to a savings for non-exemplars.⁶ But, as we saw in Experiment 3, this was not the case. Translation priming only occurred for exemplars. Also, this analysis may be too strong in its suggestion that the similarity of reference is not relevant in lexical decision because we know that masked L1–L2 translation priming occurs in lexical decision (see also Experiment 4a).

As noted earlier, there is another task that yields L2–L1 priming, namely episodic recognition (Jiang & Forster, 2001). In this study, there were two phases to the experiment. In the first phase, participants were asked to study a list of L1 words. In the second phase, participants had to indicate as quickly as they could with a button press response whether or not they had seen the L1 target during the study phase. Using this speeded “old–new” task, Jiang and Forster (2001) reported significant masked translation priming effects in the L2–L1 direction for old, but not new, targets. How can this effect be reconciled with the Sense Model account offered here? In an old–new task, the response is clearly determined by whether the L1 target stimulus matches the episodic record of the earlier presentation of the same word. Since semantic factors will be relevant in the encoding of the

⁶ This depends on the syllogism. Such a prediction is valid if it is of the form “If X is not a member of category, then it is likely that Y is also not a member, if X and Y are very similar objects.”

stimulus during the study phase, it is relevant to ask whether the episodic record includes all of the senses associated with the to-be-remembered (TBR) word, or just the dominant sense. Previous work investigating how words are encoded into episodic memory has indicated support for an encoding specificity principle (Tulving & Thomson, 1973), which assumes that only a single sense of the TBR word is encoded during the study phase (this being determined by context). In an uncued learning condition such as the one used by Jiang and Forster (2001), it is reasonable to assume that participants would encode only the dominant sense of the TBR word during the study phase. Since it is frequently the case that L1 and L2 translation equivalents share the dominant sense of the L1 word, this means that the masked L2 prime could have facilitated decision latencies on the L1 target. In the test phase of the experiment, decisions presumably are not made on the lexical status of the targets, but, rather, on whether or not the target can cue a recollection of a TBR word. If, as Tulving and Thomson (1973) have argued, the TBR word is limited in its meaning to a particular sense, which we have argued is frequently the very sense that the L2 prime is capable of activating, then we should not be surprised to observe priming in this task. Hence, the Sense Model is able to provide an explanation for the episodic priming results reported by Jiang and Forster (2001).

There are two further issues that deserve attention before moving on to the final section. The first of these is the subset model proposed by Dufour and Kroll (1995), which is potentially at odds with the assumptions of the Sense Model. These researchers found that less fluent bilinguals were able to categorize L2 targets faster when the category (e.g., COLOR) was given in L2 compared to when it was given in L1. Dufour and Kroll (1995) argued that during the earliest stages of L2 acquisition, L2 words are associated directly with a small but well-defined set of conceptual representations, essentially a subset of the category knowledge in L1. Following from this, they suggested that when both cues and targets were given in L2, only exemplars belonging to this subset of conceptual representations were activated. Consequently, there was less interference from related concepts relative to when L1 cues were used and, hence, facilitation. According to their model, L1 forms are associated with a wide range of competing concepts, all of which become active whenever the L1 form is encountered. The consequence of this widespread activation, according to their argument, is that it "...may actually inhibit the retrieval of concepts that are (otherwise) accessible from L2, because a large number of concepts that are unknown in L2 will also be activated" (p. 176). The findings reported in the present article directly contradict this conclusion. For example, the robust L2–L1 priming effects observed in Experiment 1 demonstrate that category cues presented in L1 do *not* prevent

successful processing of the L2 prime. Furthermore, by assuming that L1 cues *interfere* with L2 processing, it is not clear how the subset model could account for the robust priming effects frequently reported in the L1–L2 direction (see Introduction), which indicate that activation caused by the L1 prime *facilitates* L2 processing.

How then can the interference effects reported by Dufour and Kroll (1995) be reconciled with the facilitation effects reported here? We feel that the difference between the two findings may be accounted for by a "switch cost" present in the Dufour and Kroll study but not in ours (cf. Meuter & Allport, 1999). In both studies, participants were presented with a category exemplar (e.g., ANIMAL), which, importantly, may also serve as a language cue. In the Dufour and Kroll study, participants may have been slowest overall to categorize L2 targets when category cues were in L1 because they had to suppress their just-activated L1 in order to categorize the L2 target. This would not have been necessary when both category cue and target were given in L2. In our study, no "switch cost" was observed because (a) both category cue and target were in L1 and (b) because participants were unaware of the L2 prime. Crucially, though, the findings of the present study, which confirm those reported earlier by Grainger and Frenck-Mestre (1998), make it clear that category cues presented in L1 do not interfere with processing of the L2 prime.

One further issue that requires some discussion concerns the fact that we have used an interpolated mask between the prime and target in each of the semantic categorization experiments (as have other investigators, e.g., Bueno & Frenck-Mestre, 2002; Frenck-Mestre & Bueno, 1999; Grainger & Frenck-Mestre, 1998), whereas this is not normally done in lexical decision experiments (e.g., Experiments 4a and b). This raises the possibility that it is the interpolated mask that produces symmetry of priming, not the task. However, this is not the case. In Experiment 2, the task was lexical decision with L2–L1 translation priming, and an interpolated mask was used (as in Experiment 1), yet no priming was obtained. Also, Jiang (1999) attempted to obtain L2–L1 priming in a lexical decision task by interpolating a mask, but did not succeed.

Nevertheless, it is worth considering why such a procedure is used in the semantic categorization task (cf. Bueno & Frenck-Mestre, 2002; Frenck-Mestre & Bueno, 1999; Grainger & Frenck-Mestre, 1998; and the experiments reported here). The most obvious possibility is that reliable priming in this task *requires* the insertion of a mask. Some support for this possibility is provided by the fact that a pilot version of Experiment 5 accidentally omitted the mask, and failed to produce any priming. Clearly, further work is required to establish whether the mask is really required, but it is nevertheless of interest to consider why it might be relevant. The original purpose of including a mask was

to give the processing of L2 primes more of a headstart over the L1 targets in a lexical decision task (e.g., Jiang, 1999). The fact that this procedure failed to produce priming suggested that the mask was irrelevant, but we can now say that this failure was probably due to the nature of the lexical decision task. Thus, it might be the case that in a semantic categorization task, extra time for the processing of the prime is required when the overlap between the prime and target is purely semantic or conceptual in nature.

Within-L2 masked repetition priming

The final point that we consider here has to do with within-L2 masked repetition priming, which has been observed to be quite robust (e.g., Gollan et al., 1997; Jiang, 1999; and Experiment 2 in the present study). One of the strengths of the Sense Model lies in its ability to provide a single account of the masked translation priming asymmetry and within-L2 repetition priming. According to this account, there are no weak connections and so the masked L2 stimulus serves to activate its corresponding L2 lexical representation, which serves to activate its lexical semantic representation. As we have already argued, when the targets are in L1, the proportion of the target's lexical semantic representation activated by the L2 prime may be insufficient to produce a noticeable priming effect. However, when the targets are repetitions of the L2 prime, the target, by virtue of sharing the same orthographic and lexical semantic representations, will already have been preactivated by the prime. This should lead to very robust priming effects—and it does.

Conclusion

The results reported in the present paper provide evidence of a task difference in masked translation priming in the L2–L1 direction, with priming being observed in semantic categorization but not lexical decision. This pat-

tern of findings was interpreted to suggest that previous accounts of the asymmetry, which have proposed that weak L2 form–meaning connections constitute a limiting factor in translation priming, cannot be correct. Alternatively, we have proposed the Sense Model. This model attributes the translation priming asymmetry in lexical decision to the natural consequence of a representational asymmetry between L1 and L2 lexical semantic representations and claims that semantic categorization introduces a filtering process, which serves to eliminate the representational asymmetry. We tested these claims directly by recreating the assumed representational asymmetry with within-language word pairs (e.g., “skull–HEAD”), in which one had many senses (HEAD) and one had only one or two senses (skull). Using these word pairs, we reproduced the same pattern of priming found between translation equivalents. That is, many-sense words primed few-sense words in lexical decision and semantic categorization, but the same word pairs in the few-sense to many-sense direction produced priming only in semantic categorization, not lexical decision. These results, paralleling directly those observed in translation priming, provide strong support for the assumptions of the Sense Model.

Acknowledgments

The authors gratefully acknowledge support for this research from a number of sources, including the Faculty Small Grants Program (Office of the Vice President for Research and Graduate Studies, in conjunction with the University of Arizona Foundation), Grant NIDCD DC-01409 (as part of the National Center for Neurogenic Communication Disorders, University of Arizona), and the Cognitive Science Program, University of Arizona. We would also like to thank Jonathan Grainger, Judith Kroll, Tamar Gollan, Merrill Garrett, and Albert Costa as well as two anonymous reviewers for helpful comments and suggestions on an earlier draft of this paper.

Appendix A Experimental items used in Experiment 1^a

Part of a building		A relative		Part of the human body		A kind of metal	
roof	屋根	aunt	叔母	eye	目	iron	鉄
wall	壁	uncle	叔父	finger	指	copper	銅
window	窓	father	父	foot	足	gold	金
floor	床	mother	母	nose	鼻	silver	銀
room	部屋	grandfather	祖父	ear	耳	steel	鋼鉄
basement	地下室	husband	夫	hand	手	mercury	水銀
hall	廊下	wife	妻	mouth	口	zinc	亜鉛
stair	階段	son	息子	stomach	胃	bronze	青銅
attic	屋根裏	daughter	娘	shoulder	肩	lead	鉛
bathroom	浴室	grandmother	祖母	neck	首	potassium	カリウム

Appendix A (continued)

Color		Unit of time		Reading material		A science	
black	黒	minute	分	Magazine	雑誌	chemistry	化学
green	緑	second	秒	Book	本	physics	物理学
red	赤	year	年	newspaper	新聞	psychology	心理学
yellow	黄色	day	日	Novel	小説	biology	生物学
white	白色	month	月	thesis	論文	astronomy	天文学
maroon	栗色	week	週	Letter	手紙	mathematics	数学
brown	茶色	century	世紀	Essay	随筆	anatomy	解剖学
gray	灰色	semester	学期	Poem	詩	philosophy	哲学
purple	紫	eternity	永遠	Dictionary	辞書	geology	地質学
blue	青	hour	時間	critique	論評	medicine	医学
Animal		Profession		Insect			
cat	猫	doctor	医者	ant	アリ		
cow	牛	lawyer	弁護士	bee	ハチ		
dog	犬	dentist	歯医者	grasshopper	バッタ		
horse	馬	professor	教授	spider	クモ		
pig	豚	nurse	看護婦	cockroach	ゴキブリ		
elephant	象	secretary	秘書	termite	シロアリ		
sheep	羊	psychologist	心理学者	cricket	コオロギ		
goat	山羊	banker	銀行家	flea	ノミ		
wolf	狼	farmer	農家	beetle	カブトムシ		
bear	熊	teacher	先生	fly	ハエ		

^aCritical items are shaded grey.

Appendix B. Experimental items used in Experiment 4

One sense	Many sense	# of senses for many-sense words
oxygen	AIR	20
moist	WET	8
movie	FILM	7
sofa	CHAIR	7
surgeon	DOCTOR	7
oar	PADDLE	10
tomb	GRAVE	8
violin	BASS	9
clergy	MINISTER	6
elderly	AGED	10
wealthy	RICH	12
coyote	FOX	9
lantern	LIGHT	30
huge	LARGE	12
lawn	GRASS	10
wheat	GRAIN	11
lung	HEART	10
cottage	HOUSE	13
dairy	MILK	7
poverty	POOR	9
saucepan	POT	11
fist	HAND	15
pistol	GUN	8
prince	KING	6
beer	DRINK	11
fabric	MATERIAL	12
lawyer	JUDGE	7
tiny	SMALL	17

Appendix B (continued)

One sense	Many sense	# of senses for many-sense words
pebble	STONE	10
marina	DOCK	11
wrist	ARM	8
ankle	FOOT	12
lengthy	LONG	15
puppy	DOG	7
kitten	CAT	9
gale	WIND	16
suitcase	BAG	15
creek	STREAM	10
mosquito	FLY	20
rapidly	FAST	15
	Mean	11.25

Appendix C. Experimental items used in Experiment 5

Many-sense words	# of senses	Few-sense words	# of senses
dress	19	gown	4
slacks	12	trousers	1
skirt	7	kilt	1
coat	6	parka	1
jacket	7	windbreaker	1
chair	6	recliner	1
table	7	desk	1

Appendix C (continued)

Many-sense words	# of senses	Few-sense words	# of senses
couch	4	sofa	1
bed	9	futon	1
dresser	5	cupboard	1
dove	8	pigeon	1
eagle	5	osprey	1
crow	9	blackbird	2
canary	5	sparrow	1
saw	25	lathe	1
hammer	10	mallet	3
drill	9	awl	1
file	9	chisel	1
wrench	7	pliers	1
red	9	crimson	2
blue	19	lavender	2
green	15	teal	2
purple	8	violet	3
black	20	ebony	2
dog	7	puppy	2
cat	9	kitten	2
horse	7	zebra	1
cow	4	bison	1
wolf	6	coyote	1
pig	9	boar	2
bear	15	grizzly	1
knife	4	cleaver	1
spoon	5	ladle	2
cup	11	goblet	1
pan	7	skillet	1
plate	15	platter	1
head	40	skull	1
legs	9	thigh	2
hand	15	wrist	1
foot	12	ankle	1
eye	6	retina	1
nose	13	nostril	1

References

- Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access. The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 340–357.
- Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology*, 80, 1–46.
- Besner, D., & Smith, M. C. (1992). Models of visual word recognition: When obscuring the stimulus yields a clearer view. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 468–482.
- Bodner, G., & Masson, M. (2001). Prime validity affects masked repetition priming: Evidence for an episodic resource account of priming. *Journal of Memory and Language*, 45, 616–647.
- Bueno, S., & Frenck-Mestre, C. (2002). Rapid activation of the lexicon: A further investigation with behavioral and computational results. *Brain and Language, Special Issue: Mental lexicon II*, 81(1–3), 120–130.
- Catchpole, J. (1987). *The locus of the word frequency effect*. Unpublished honors thesis, Department of Psychology, Monash University.
- Cree, G., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science*, 23(3), 371–414.
- Damian, M. F. (2001). Congruity effects evoked by subliminally presented primes: Automaticity rather than semantic processing. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 154–165.
- Davis, C.W., Kim, J., & Sanchez-Casas, R.M. (2003). Masked priming across languages: An insight into bilingual lexical processing. In S. Kinoshita & S. Lupker (Eds.), *Masked priming: State of the art*.
- de Groot, A. M. B. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 1001–1018.
- de Groot, A. M. B. (1993). Word-type effects in bilingual processing tasks: Support for a mixed-representational system. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 27–51). Amsterdam: John Benjamins.
- de Groot, A. M. B., & Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language*, 30, 90–123.
- 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.
- Dufour, R., & Kroll, J. F. (1995). Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation. *Memory & Cognition*, 23, 166–180.
- Fellbaum, C. (Ed.). (1998). *WordNet: an electronic lexical database*. Cambridge, MA: MIT Press.
- Finkbeiner, M. (2002). Bilingual lexical memory: Towards a psycholinguistic model of adult L2 lexical acquisition, representation, and processing. (Doctoral dissertation, University of Arizona, 2002). *Dissertation Abstracts International*, 63 (12), 6114B. (UMI No. AAT 3073282).
- Finkbeiner, M. (in press). Task-dependent L2–L1 translation priming: An investigation of the separate memory systems account. In *Proceedings of the 4th international symposium on bilingualism*. Somerville, MA: Cascadilla Press.
- Forster, K. I. (2000). The potential for experimenter bias effects in word recognition experiments. *Memory & Cognition*, 28, 1109–1115.
- Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35(1), 116–124.
- Forster, K. I., & Hector, J. (2002). Cascaded vs non-cascaded models of lexical and semantic processing: The *turple* effect. *Memory & Cognition*, 30, 1106–1116.
- Forster, K.I., Mohan, K., & Hector, J. (2003). The mechanics of masked priming. In S. Kinoshita & S. Lupker (Eds.), *Masked priming: State of the art*. Psychology Press.

- Forster, K.I. (in press). Category size effects revisited: Frequency and masked priming effects in semantic categorization. *Brain and Language*.
- Frenck-Mestre, C., & Bueno, S. (1999). Semantic features and semantic categories: Differences in rapid activation of the lexicon. *Brain and Language*, *68*, 199–204.
- Frenck-Mestre, C., & Prince, P. (1997). Second language autonomy. *Journal of Memory and Language*, *37*, 481–501.
- Gollan, T., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew–English bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23*, 1122–1139.
- Grainger, J., & Frenck-Mestre, C. (1998). Masked priming by translation equivalents in proficient bilinguals. *Language and Cognitive Processes*, *13*, 601–623.
- Jiang, N. (1999). Testing processing explanations for the asymmetry in masked cross-language priming. *Bilingualism: Language and Cognition*, *2*, 59–75.
- Jiang, N., & Forster, K. I. (2001). Cross-language priming asymmetries in lexical decision and episodic recognition. *Journal of Memory and Language*, *44*, 32–51.
- Keatly, C. W., Spinks, J. A., & De Gelder, B. (1994). Asymmetrical cross-language priming effects. *Memory & Cognition*, *22*, 70–84.
- Kroll, J. F., & de Groot, A. M. B. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 169–199). Mahwah, NJ: Erlbaum.
- 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.
- Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*, 25–40.
- Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective view of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.
- Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. *Quarterly Journal of Experimental Psychology A*, *55*, 1339–1362.
- 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.
- Sanchez-Casas, R. M., Davis, C. W., & Garcia-Albea, J. E. (1992). Bilingual lexical processing: Exploring the cognate/noncognate distinction. *European Journal of Cognitive Psychology*, *4*, 293–310.
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 359–380.
- Uyeda, C., & Mandler, G. (1980). Prototypicality norms for 28 semantic categories. *Behavior Research Methods, Instruments & Computers*, *12*, 587–595.
- Williams, J. N. (1994). The relationship between word meanings in the first and second language: Evidence for a common, but restricted, semantic code. *European Journal of Cognitive Psychology*, *6*, 195–220.