4
The Nature of the Bilingual Lexicon: Experiments with the Masked Priming Paradigm

Kenneth I. Forster and Nan Jiang

The study of bilingualism provides a special opportunity for psycholinguists and neuroscientists to study the organization of memory, in this case, memory for lexical items. The kinds of questions that can be raised are many and varied. Is there a separate mental lexicon for each language, such that lexical items are independently associated with their appropriate semantic and syntactic representations, or is there a common lexicon, in which items from different languages are effectively intermingled? Are the two lexicons associated with different functional neural subsystems? Is there a bilingual “switch” which turns one language on, and other languages off? Is a late-acquired language (L2) represented in the same way as the native, first-learned language (L1)?

The evidence we will be discussing is drawn from studies dealing with the recognition of the printed word. It argues for a rather unexpected conclusion, namely, that the lexicons for each language are entirely distinct, independently and separately accessed. Not only this, but for late learners at least, they appear to be represented in entirely different processing systems. Specifically, the evidence seems to suggest that the L1 lexicon is represented within the language module (the part of the brain devoted to language), while the L2 lexicon is not.

To explain how this conclusion was reached, we need to consider the methods used in our experiments. The basic technique uses the lexical decision task. This technique involves presenting a sequence of letters on a computer display, and asking the subject of the experiment to decide as quickly as possible whether it is a familiar word or not. Half of the items shown to the subject are actual words, and half are not words at all, although they are perfectly good potential words (e.g., blink). This task raises an interesting problem, first addressed by Norman (1969), who referred to it as the manteness problem. Norman was interested in the question of how we could decide so rapidly that manteness was not a word in English. The argument is that this decision is made so rapidly that we could not have searched through a large inventory of lexical forms (more than 60,000 items for the average college-educated individual), but instead, we must have some direct procedure for testing whether the form exists. This involves some kind of content-addressable memory, such as that involved in a library catalogue, or an actual dictionary. To see whether manteness is a word, it is not necessary to scan the entire dictionary — all that is required is to find the place in the dictionary where this word would be stored, if it was a word.

The time it takes the subject to make this judgment provides us with our dependent variable, which is called the lexical decision time. By observing what kinds of experimental manipulations make a difference to lexical decision time, we can make inferences about the nature of the neural machinery that is responsible for retrieving the relevant contents of lexical memory. The particular manipulation of interest here involves presenting another word just prior to the target word. If this word bears some kind of similarity to the target, we may observe that the reaction time to the target is decreased. This would be termed a positive priming effect, and the assumption is that the processing carried out on the prime (the first word) alters the nature of the processing carried out on the target, either by making the whole process faster, or by eliminating some steps in the computation. It is also possible that the prime may increase the time taken to recognize the target. This would be termed an inhibitory, or negative priming effect.

Priming experiments have been carried out by presenting words that have a similar meaning to the target, e.g., boy-prince, or are associatively related to it, e.g., boy-girl (for a review, see Neely, 1991). Also, priming has been studied where the relationship is one of form rather than meaning. In this case, the prime resembles the target either in orthography (e.g., laughter-laughter) or in phonology (e.g., ate-eat). By far the strongest form of priming occurs when the prime and target are in fact the same word, i.e., repetition priming. When it comes to the study of bilingualism, the interesting case is when the prime is a translation equivalent term in another language known to the individual. So if we were studying a Spanish-English bilingual, we might present the target woman for lexical decision, and precede it with the translation prime in Spanish, mujer. By comparing the mean lexical decision time for this target word with that obtained for the same target word preceded by a completely unrelated Spanish word, we can determine whether a translation prime has any effect.

Why should a translation prime have any effect? The answer concerns the types of representations involved. If two words have overlapping representations (i.e., their representations have common elements), then there is a chance that some of the neural elements activated by the target will already have been activated by the prime. This leads to a much more rapid activation of the target representation. However, one of the problems that has plagued research in the priming area is that the perceived relationship between the prime and target might influence the time taken to make a decision about the target. That is, the prime may have no effect at all on the time taken to recognize the target, but the fact that the two words have some kind of congruent relationship may alter the way in which a decision about the target is reached. As an example, consider the case in which the prime is the word blanket, and the target is the word pillow, but the subject is momentarily uncertain whether it was pillow or willow, or maybe even willow (i.e. a noun). However, given that the prime was the word blanket, the subject may reason that the target most likely was in fact pillow, since the probability of misidentifying the target as a word that happens by accident to be related to blanket must be very low. This reduc-
tion in uncertainty leads to a faster decision, although no priming has actually taken place.

There are other ways in which the faster response to the target might be explained. It could be that the subject anticipates what the target will be, and the decision time is influenced by whether the subject's anticipation was correct or incorrect. This is a real problem in a bilingual priming experiment, since subjects very quickly realize that many of the target words are translations of the preceding prime, and if they actively translate the prime, the obtained priming effect could be due to correctly anticipating the target. Or it could be that subjects notice post hoc that the target word was a translation of the prime, and this congruence facilitates the decision.

Obviously, it would be desirable to design a priming paradigm in which the priming effects were independent of whether subjects were able to anticipate the target, or whether the prime and target were perceived to be congruent or not. One way in which this has been done in the past is to decrease the time interval between the onset of the prime and the target. For example, if the target appears only 250 msec. after the prime, it seems less unlikely that subjects could have generated any kind of expectation about the nature of the target in the time available. Priming effects obtained under these conditions are often thought to be completely automatic effects. However, this procedure does nothing to eliminate post hoc congruence effects.

The procedure we adopted was to use a masked priming paradigm (Forster & Davis, 1984). In this procedure, the prime is very briefly presented (50 msec.), and is followed immediately by the target, which overwrites the prime. The prime is always physically different from the target. In Roman scripts this is achieved by having the prime in lower-case letters, and the target in upper-case letters. In other scripts where there is no counterpart to case, different fonts and different sizes are used. In addition, the prime is preceded by a forward mask, and the masking effect from this stimulus, combined with the backward masking effect of the target stimulus, renders the prime virtually invisible for the vast majority of subjects. We reasoned that if the subject was quite unaware of the existence of the prime, then it could not be used to form an expectation about the target, and could there be any post hoc appreciation of the congruence between the prime and target.

This technique has been used quite extensively in bilingual research. Diane Bradley and several of her students and colleagues have investigated cross-language priming in several languages (Garcia-Alba, Bradley, Sánchez-Cassas & Forster, 1985; Milkskoed, 1986; C. Lanza, 1988). The upshot of this work suggests that translation priming is restricted to cognates, i.e., words with similar phonology and orthography. For example, Sánchez-Cassas, Davis and Garcia-Alba (1992) tested Spanish-English bilinguals in a semantic categorization task where the prime was either the same word as the target, a translation equivalent word, or an unrelated word. The priming direction was from L2 to L1. It was found that cognate translations (e.g., rich–rich) produced significant priming effects, which were comparable to identical word pairs. However, no priming effects were found for noncognate translation pairs.

Other investigators have also found stronger priming for cognates but recent work suggests that this result may depend on the type of script. Gollan, Forster and Frost (1997) found equally strong priming for both cognates and noncognates with Hebrew-English bilinguals. These two languages have quite different scripts, whereas previously studied languages use the same, or similar, scripts. Gollan et al. suggested that when the prime language has the same script as the target language, access is attempted initially in the lexicon appropriate to the target language. Thus, if the experiment involves Spanish primes and English targets, the subject (being quite unaware that anything other than English words are being displayed) might be set to process only English words. This would mean that an attempt is made to first access the prime in the English lexicon, which would obviously fail (except perhaps in the case of cognates, which might be jointly listed). This does not mean that the masked Spanish prime would go unrecognized. It is a common occurrence in a monolingual lexical decision experiment for a subject to report that some of the nonwords were actually words in another language. This means that both lexicons eventually are accessed. However, priority might still be assigned to the target language, so that the initial attempt to access the prime is made in the English lexicon, and if this fails, a subsequent attempt is made in the Spanish lexicon. This latter attempt, however, would be seriously delayed, which would reduce the possibility of priming.

However, if the primes are in Hebrew and the targets are in English, it would be strange to suggest that an attempt is made to access the Hebrew prime in the English lexicon, since the characters are completely different. It seems far less likely that the Hebrew characters would automatically engage the Hebrew lexicon, and the Roman characters would automatically engage the English lexicon, and that this would override any expectation. Thus we should expect the strongest translation priming effects in two languages with different scripts. Thus, for Hebrew-English bilinguals, both cognates and noncognates show strong priming effects, whereas for Spanish-English bilinguals, only cognates show any effects.

Support for this interpretation has been provided recently by Jiang (1999), who confirmed the existence of strong translation priming for noncognates in a Chinese-English experiment using a lexical decision task. This effect is what the script hypothesis predicts, since there is no overlap at all between the two scripts. An interesting exception to this trend has been reported by Grainger and Frenck-Mestre (1998), who found priming for French and English. However, they used a semantic categorization task, in which subjects are required to categorize the target word according to its semantic properties (e.g., is it the name of an animal?). When a lexical decision task was used, no significant priming effects were obtained. Just how this difference in tasks is relevant is not at all clear.

But for present purposes, there is a more important property of the cross-language priming effect, namely, its asymmetry. In both the Hebrew-English and the Chinese-English studies, priming was observed only if the prime was in the subject's dominant language (L1), and the target was in the subordinate language (L2). No priming at all was observed in the reverse direction (L2 to L1). One obvious consideration is that the subject may be unable to process L2 primes when they are presented so briefly. However, this cannot be the explanation for the asymmetry, since L2 primes are quite effective with L2 targets. For example, in the Gollan study, English-dominant bilinguals showed no priming with Hebrew primes and English targets, but with Hebrew targets and Hebrew primes, strong priming was obtained. The same was true for Chinese-dominant bilinguals in the Jiang study. English primes were ineffective with Chinese targets, but were quite effective with English targets.

One might attempt to explain these results by appeal to the relative speed of processing of words in L1 and L2. An L2 prime might be processed too slowly with re-
spect to an L1 target to have any effect. In essence, the process initiated by the L1 target "overrides" the process initiated by the L2 prime. However, when the target is also in L2, the prime and target are processed at the same rate, and hence the prime process has a chance to reach completion before the target process. This explanation suggests that L2-L1 priming might be possible if the target was delayed, giving more time for the slower L2 prime process to complete. This hypothesis was put to the test by Jiang (1999), but the outcome was not encouraging. No L2-L1 priming was observed despite a delay of 250 msec., which should have been long enough for the prime to be processed.

This is not the first time that such an asymmetry has been reported. Similar effects can also be found in most previous non-masked cross-language priming studies where priming effects have been consistently greater from L1 to L2 than from L2 to L1 (Jin, 1990; Altarriba, 1992; Keatley et al., 1994). What lies behind this asymmetry? One possible account, the working mode hypothesis, is based on the distinction between monolingual and bilingual working modes (Soares & Grosjean, 1984; Grosjean, 1992). In the former case only the language in use is activated, in the latter, both languages are in an active state. Which mode a bilingual speaker works in depends on several factors. One of them is the immediate working language. For a non-balanced bilingual speaker, it may be easier to suppress the less proficient L2 and keep it at a minimum activation level (Green, 1985, 1998). Thus, when the L1 is the immediate working language and no use of L2 is required, the L2 speaker may effectively work in a monolingual L1 mode. On the other hand, when L2 is the working language, L1 may also be active. In a cross-language priming experiment where the prime is masked, the subjects' working mode may be determined by the dominance of the target word. If the target language is L1, subjects will be working in a monolingual L1 mode. As a result, the L2 may be deactivated or suppressed, and hence the L2 prime is less likely to be processed. When the target language is L2, however, subjects will be working in a bilingual mode, and hence the L2 prime will be processed and will affect the processing of the L2 target. To test this hypothesis convincingly, it would be necessary to allow independent indicators of which mode the subject is in, without such an indicator, all that can be done is to devise situations that ought to push the subject in one direction or another. Jiang (1999) made an attempt in this direction by mixing the target languages instead of presenting them in a blocked manner. That is, instead of having one session consisting purely of Chinese targets, and another of English targets, Jiang randomly intermixed items so that the target language was unpredictable. This ought to push the subject towards a bilingual mode, but Jiang failed to find any evidence of L2-L1 priming under these conditions. Another method currently being explored in our laboratory by Adam Dudis involves presenting an L2 word prior to the masked L2 prime, so that the resulting sequence is: L2 word, masked L2 prime, L1 target. From the subject's point of view, every trial consists of a word in L2 followed by a word in L1, and one might imagine that this would be enough to keep the subject in bilingual mode. However, current results suggest that this procedure has no effect. There is still no L2-L1 priming.

Kroll and Stewart (1994) suggested that the asymmetry may reflect some representational features of the bilingual memory. According to their revised hierarchical model, L2 words are related to their equivalents in L1 through both lexical links and shared conceptual representations (see Figure 3.1 in Kroll & Tokowicz, this volume). The model further proposes dual asymmetries in the strengths of these connections. Connections with conceptual representations are stronger for L1 words than for L2 words, and lexical connections are stronger for L2 to L1 than in the reverse direction. The hypothesis that lexical linkages between L2 words and their translation equivalents in L1 play an active role in bilingual language processing is supported by two lines of evidence from studies by Kroll and her colleagues. First, both early and proficient bilingual speakers were able to translate faster from L2 to L1 than from L1 to L2 (Kroll & Stewart, 1994; Sholl, Sanakaranarayanan & Kroll, 1995). This translation asymmetry was interpreted as reflecting the employment of different links in performing translation in two different directions. The direct lexical links are utilized in L2-L1 translation, which produce a faster performance than L1-L2 translation in which conceptual links are involved. The second line of evidence comes from the finding that conceptual factors affecting translation from L1 to L2, but not L2 to L1 (Kroll & Stewart, 1994; Sholl et al., 1993). It was found in Sholl et al., for example, that a task assumed to involve conceptual representations (picture naming) facilitated later translation from L1 to L2, but not from L2 to L1.

The implications of this model for cross-language priming are twofold. If we attempt to explain cross-language priming purely in terms of overlapping conceptual representations, then it is difficult to explain why symmetric priming effects are not found. However, since the connections between L2 words and conceptual representations are assumed to be weaker than for L1 words, it might be suggested that under masked priming conditions, the L2 connections are too weak (or too slow) to exert any appreciable effect, and this may explain why there is no priming at all from L2 to L1. The problem with this explanation is that it appears that the masked priming technique is not particularly sensitive to conceptual overlap. Within-language semantic priming experiments using semantically related words (e.g., animal-tiger) or strongly associated words (e.g., bread-butter) produce effects that are notoriously weak when the primes are masked (e.g., Perez & Gómez, 1997; Forst, Forster & Deutsch, 1997), whereas the L1-L2 translation priming effects are much stronger. Admittedly, translation equivalent terms are likely to be closer in meaning than the related items typically used in semantic priming studies, but this does not seem to be sufficient to account for the discrepancy.

The other possibility provided by the Kroll and Stewart model is that the direct linkages between the two lexicons might play a role in explaining the asymmetry, since these lexical linkages are themselves asymmetric, in that the L2-L1 connections are assumed to be stronger than the L1-L2 connections. If two words are strongly linked, then activating one will activate the other, and this should produce priming. However, if the linkages are stronger in the L2-L1 direction, this would suggest that priming should be stronger from L2 to L1, which is the exact opposite of what is needed. So this mechanism scarcely helps. To avoid predicting the wrong asymmetry, one would have to assume that these linkages are only activated when the subject is aware of the L2 prime, suggesting that these associations are not entirely automatic, but instead depend on the consciousness intention to translate. Thus, these linkages are invoked in a translation task, but not in masked priming.

So, unless we accept the conceptual overlap account of masked priming, it appears that we are left without any wholly satisfactory explanation of the asymmetry. One way to improve the situation would be to find a set of conditions that did produce L2-L1 priming, since this would provide a vital clue. However, the work reported in Jiang (1999) represents three attempts to modify conditions to facilitate L2-L1 priming, and none of them had any effect. However, one additional procedure did produce a dramatic
effect (Jiang & Forster, under revision), and we now focus on this effect, and its implications.

The idea for this experiment came from work carried out by Bradley (1991). She investigated the so-called city-grass effect reported by Mckoon and Ratcliff (1979, 1986). This experiment challenged the traditional distinction between episodic memory (memory for events) and semantic memory drawn by Tulving (1972). Their claim was that if subjects learned to associate unrelated words such as city and grass, then this episodically established association would act just like a semantic association in producing priming in a lexical decision experiment. That is, grass would be classified as a word faster if it was primed by city. Bradley was able to show that this effect did not occur if the prime was masked. However, this demonstration was vulnerable to the criticism that an episodically established association may be too weak to survive masking. If training were continued for some unspecified period of time, then masked priming might eventually be obtained. In order to cope with this argument, Bradley needed to demonstrate that the association was strong enough to play a role in a masked situation. To do this, she adopted an unusual procedure. The stimuli remained exactly the same, but the task was changed. Instead of classifying grass as a word, subjects were asked to decide whether gras was one of the words they had learned in the previous list. This task is referred to as a speeded episodic recognition task (also known as an old-new task). Under these conditions, very strong priming effects were obtained. That is, the old-new decision was made faster when the target was primed with its associated (masked) partner in the previous list. This demonstrates that the association between city and grass was strong enough to produce a masked priming effect, providing that an episodic task was used.

The implication of this finding is that the city-grass connection was established in episodic (or non-lexical) memory, and under masked priming conditions, its existence can only be demonstrated if the task is focused on the contents of the episodic memory system, not the lexical system. On the other hand, the semantic association between doctor and nurse is established in lexical memory, and its existence can be demonstrated (albeit weakly) in a lexical task with masked priming but not in an episodic task (however, this latter point has not yet been demonstrated).

The relevance of this experiment to the bilingual case should now be obvious. Perhaps the L2-L1 lexical connections posited by Kroll and Stewart are analogous to the city-grass connections. Perhaps they exist only in episodic memory, and are therefore ineffective when the task is focused on output from the lexical system. If so, then it should be possible to obtain L2-L1 priming if an episodic recognition task is used instead of a lexical decision task.

The results of the first experiment of this type are reported in Jiang and Forster (under revision). The experiment involved a group of 26 Chinese-English bilingual speakers from mainland China studying at the University of Arizona at the time of testing, who would be best described as late learners of English. Most of them had studied English in formal settings for at least eight years, and the number of years spent in the U.S. ranged from less than a year to eight years. The episodic recognition task was divided into two phases, the study phase and the test phase. During the study phase, 32 Chinese two-character bisyllabic words were presented one at a time, and subjects were told to remember them. The test phase involved the presentation of the same 32 words ("old" items) intermixed with 32 "new" words not seen previously. Each trial consisted of the following sequence: a forward pattern mask for 500 msec., an English prime for 50 msec., a pattern mask for 150 msec., and then the target Chinese word for 500 msec. Subjects had to classify each target word as "old" or "new" as rapidly as possible. For half the trials, the English prime was a translation equivalent term for the target, and half the time it was an unrelated English word. This entire procedure was repeated twice, but only the data from the second test phase were analyzed, since the error rates on the first test were too high to yield meaningful estimates of reaction times.

The data from the test phase showed an L2-L1 priming effect for "old" items. That is, the Chinese targets were classified as "old" faster (by an average of 29 msec.) when they were preceded by a masked English translation than when they were preceded by a masked unrelated English word. This effect was significant in both subject and item analyses. However, for "new" items, the effect of the prime was not significant, and was slightly in the wrong direction. Thus it appears that L2-L1 priming does occur in an episodic recognition task, but only for "old" items. This differential effect for "old" and "new" items is an important result, since if both items showed priming, it could be argued that it was some aspect of the task itself that produced priming. For example, the episodic task might be more difficult and require subjects to pay more attention to the input, and this is why priming occurs. But the absence of priming for "new" items suggests that this is not the case. Instead, it appears that the critical factor is the attempt to retrieve episodic memory traces rather than lexical memory traces. Finally, as a check on earlier results, these subjects were also run in a lexical decision task using half of the items used in the episodic task. The data confirmed that there was no significant L2-L1 priming in lexical decision, the effect being only +8 msec.

On the face of it, these results may not seem particularly intuitive. In the episodic task, subjects are attempting to retrieve information about whether a particular Chinese word had been shown to them before, and it is hard to see how the prior masked exposure of its English translation could possibly assist this task, especially since the same procedure is of no benefit in a lexical decision task. However, if we assume that lexical items in L2 are represented in episodic memory, and gain access to conceptual representations only through episodically established associations with lexical items in L1, then it is reasonable to suppose that these associations might only become active when the subject's processing resources are largely devoted to the task of episodic retrieval.

This much seems reasonable, but how does this produce priming? The answer requires a brief description of an earlier series of experiments (Forster, 1985). This study concerned the issue of lexicalization. How quickly do new words become part of the lexicon, and how can we tell? One possible indicator is the existence of masked priming, since only lexical items normally show strong masked priming effects. Thus one might expect novel words to show no masked priming initially, but as they become more familiar, the strength of the priming effect would gradually increase. This proved not to be the case. Subjects were given new English words to learn, which consisted of obsolete words such as holomonth, and told to treat these items as words in a subsequent lexical decision task. Surprisingly, these newly acquired words showed strong masked priming effects immediately after learning. Further investigation revealed that masked priming effects could also be obtained for nonwords in a conventional episodic recognition task, indicating that episodic memory traces could be primed in much the same way as lexical memory traces. This means that the reason that holomonth showed such strong priming could
be that lexical decisions concerning such a word required the retrieval of an episodic trace, not a lexical trace. Given this information, we can explain priming in the present bilingual situation by proposing that in an episodic recognition task, the L2 prime activates the episodic representation of the L1 target established during the study phase.

To illustrate the logic of this argument, let us represent the lexical entry of a word as M, and its episodic representation as M*. By hypothesis, words in L2 have only episodic representations (represented as M2*), which are linked to the episodic representations of L1 words (represented as M1*), which in turn are connected to lexical representations (M). Thus,

1. L2 prime → M2* → M1* → M

Sequence (1) suggests that L2 words ought to prime words in L1, but this sequence is not entirely automatic. We assume that the entire sequence is run off only when subjects are aware that they are required to interpret words in L2. This is the case if the subject is aware of the L2 prime, but not if it is masked. In the latter case, the sequence is broken at some point prior to the activation of M1, which is a requirement for L2-L1 priming in a lexical decision task. This is shown in sequence (2):

2. Masked L2 prime → M2* → M1*

However, when the task is explicitly episodic in nature, and the target is a word in L1, then the decision will be made on activation in M1* (not M), and hence priming from a masked L2 prime can occur.

This explanation leaves some loose ends. For example, why is there no priming for the "new" items? The most plausible answer is that the target items do not have a task-relevant episodic trace, and hence there is no prime to be primed. (In these cases, the correct response is "I do not know." The timing of such a decision might be controlled by a deadline. If no evidence of an episodic trace has been discovered after a certain time has elapsed, then a "No" decision is made. It is not surprising that the masked L2 prime has no effect on such a decision.

Another loose end concerns the existence of L2-L2 masked priming effects in lexical decision. Both Gollan et al. (1997) and Jiang (1998) demonstrated strong L2-L2 priming effects in lexical decision, despite the absence of any L2-L1 priming. How is this possible? If a masked L2 prime never activates any lexical representations, as in sequence (2), the answer is that such subjects are unaware of the existence of the L2 prime, and hence are "set" to run off sequence (1) in its entirety whenever an L2 item is presented. So, if a "Yes" response in lexical decision depends on activation in M1, then priming should occur. An alternative explanation is that lexical decisions to L2 words depend on the retrieval of episodic traces, not lexical traces, and these will always be activated whether the prime is masked or not. Another way to say this is that a lexical decision experiment with L2 targets is really equivalent to an episodic recognition task.

Subsequent experiments in this series (Jiang & Forster, under revision) went on to replicate these findings. It was also shown that merely having a study phase was not sufficient to produce priming for studied items in a lexical decision task. Of critical importance was the fact that the L2-L1 effect depended on allowing more time than usual for processing the L2 prime. The standard masked priming paradigm has a target onset occurring immediately after prime offset, so that the interval between the offsets of prime and target is the same as the prime duration (50 msec. in these experiments). When this procedure was used, no L2-L1 priming was obtained. It was only when an additional 150 msec. interval was introduced after prime offset that any effect was obtained. The final result in this series was perhaps the most intriguing of all. An obvious question to ask is whether there is any L1-L2 priming in an episodic task. Since cross-language tasks seem to involve asymmetries, one might predict another asymmetry in this case, only this would be the reverse asymmetry. That is, L1-L2 is better than L2-L1 in lexical decision, but L2-L1 is better than L1-L2 in episodic recognition.

This was indeed the outcome. When subjects were given English words to remember, a masked Chinese translation was of no assistance in making the old/new decision. The effect here was a +6 msec. for "old" items, and +13 msec. for "new" items. Neither effect appeared significant. However, when the same subjects were tested with the same items in a lexical decision task, there was a very strong +40 msec. effect.

While it is pleasing to observe that there is a certain symmetry to the asymmetries, one might wonder whether there is any other reason to expect this result. Kroll and Stewart (1994) postulate lexical links from L1 to L2 as well as from L2 to L1. These lexical links would also be episodic in nature (our assumption, not theirs), and so there ought to be priming in an episodic task. However, Kroll and Stewart also propose that the L1-L2 links would be much weaker than the L2-L1 links, and hence it could be argued that they are not strong enough to support priming. The weaker strength of these connections results from the fact that in normal reading situations, a bilingual would seldom need to translate a word in L1 into its L2 equivalent. But when a word in L2 is encountered, it must always be translated into L1.

Remaining Questions

As noted earlier, Grainger and Ferenczi (1998) obtained L2-L1 priming when a semantic categorization task was used, but not when a lexical decision task was used. In this type of task, the subject is asked whether the target word refers to an entity that belongs to some conceptual category, e.g., animals. This finding raises obvious problems for the episodic interpretation, since we would have to explain why a semantic categorization task should activate episodic links. One would think that this task should produce a focus on lexical semantics, and that non-lexical factors would play even less of a role than they do in lexical decision. Of course, this might depend on the nature of the category. It might be that some categories require something very much

---

1 A similar result was obtained in the homophone experiment (Forster, 1983).

2 This would normally make the prime quite visible, and it was for this reason that a post-prime mask was introduced.

3 No mask was used between the prime and target. The interval between the onset of the two stimuli was 50 msec.

4 It should be noted that Sanchez-Casas et al. (1992) also used a semantic categorization task in an effort to amplify L2-L1 non-semantic priming, but no priming was found.
akin to episodic retrieval in order for rapid classification to occur. An obvious example would be a category like “made of metal”. Is it part of the core meaning of the word dagger that the objects it refers to are usually made of metal? Or is this a contingent property of daggers? If we have to search to find instances of daggers and then determine whether they are made of metal, then the task might not be so different from an old-new task.

Another possibility, of course, is that all sorts of tasks may show L2-L1 priming, all except lexical decision. This would raise a completely different set of problems. The episodic interpretation would have to be abandoned, and we would then have to seek reasons why the lexical decision task should suppress L2-L1 priming, but not L1-L2 priming. This is not an attractive prospect.

The obvious next question to ask is whether these findings also apply to highly proficient bilinguals. A translation strategy for L2 is reasonable to assume for a late learner, but surely not for bilinguals who have spoken L2 for many years and use it on a daily basis? One might expect that such individuals would show perfect symmetry in their priming, or perhaps no cross-language priming at all, on the grounds that L1 and L2 access separate lexicons, with no cross-language linkages between lexical items. In this respect, it is interesting to note that many of the Hebrew-dominant bilinguals studied in Gollan et al. (1997) were thought to be highly fluent speakers of English, yet they showed very little priming from English to Hebrew. Currently, this question is being investigated in our laboratory by Adam Dudiski, who is carrying out intensive testing with individual subjects, rather than collapsing data across a number of individuals. Although this project is incomplete, the available data suggest that highly proficient Chinese-English bilinguals also show the same asymmetry in lexical decision, even though their proficiency in L2 is so high that they are considered native speakers by their peers.

This result makes it very hard to argue that cross-language priming is conceptually mediated, at least in the masked case. For a less proficient bilingual, it is at least possible to argue that the connections between L2 and conceptual representations are too weak to survive masking, but to make the same argument for highly proficient bilinguals would rob the explanation of any force. There is an alternative view that might make sense of this result. This view assumes that once any language takes total possession of the language system, it competes with any would-be successor, relegating it to cortical areas outside the classic language areas. This view implies that priming asymmetries would remain no matter how fluent the speaker was in L2, provided that L2 was learned after L1, and provided that both languages were in constant use.

In this connection, it is interesting to note that brain imaging studies of bilinguals show striking differences in the cortical areas associated with L1 and L2 (e.g. Kim, Kelken, Lee & Hirsch, 1997; Petani, Dehaene, Grassi, Cohen, Cappa, Dupoux, Fazio & Mehler, 1996). These results are clearly consistent with the view adopted here, but it has yet to be established whether the patterns of brain activity in L2 tasks are at all similar to the patterns typically observed in episodic recognition tasks. If such a correspondence can be established, this would provide striking support for the episodic interpretation of L2-L1 priming.

Acknowledgements

This research was supported in part by National Multipurpose Research and Training Grant DC01409 from the National Institute on Deafness and Other Communication Disorders to the National Center for Neurogenic Communication Disorders at the University of Arizona, and by a grant from the McDonnell-Pew Foundation Cognitive Neurosciences Program.