CHAPTER 4

Using priming to explore early word learning

John N. Williams and Amanda Cheung
University of Cambridge

Current views of vocabulary learning imply that second language (L2) words initially inherit the meaning of their first language (L1) translation equivalents. We describe a series of experiments that test this idea by examining semantic priming from newly learned words. The aspects of meaning that these words automatically activate were probed by manipulating the nature of the prime-target relationship. Although there were clear semantic priming effects, not all types of semantic relationship supported priming. The effects of variations in study conditions lead us to suggest an episodic view of word learning in which newly learned words do not simply inherit L1 meanings, but rather are associated with the aspects of meaning that were active at the time they were learned. We discuss the implications for teaching vocabulary, especially through the use of translation equivalents.

Introduction

Imagine that you are an English native speaker who is learning French, and that you have just been told that the word *écureuil* means "squirrel." After about 10 minutes you are asked "what does écureuil mean?" and you say "squirrel." You can also create sentences (in French) like "The écureuil climbed the tree," "The écureuil ate a nut," "An écureuil is a mammal." You appear to have learned the word *écureuil*. But in what sense have you learned it? How is this novel word actually represented in your mind? Is it represented in the same way as for a native speaker of French? You might suspect not. The native speaker of French has experienced and used this word in countless different contexts over many years, so surely their representation of it must differ from that of the person who just learned it 10 minutes ago? But in what sense is the representation different? After all, you do appear to be able to use the word correctly. But how is this new word actually wired into your mind, and, importantly, how is that affected by the way you were taught?

One prevalent view is that newly learned words in a second language (L2) are simply wired to their first language (L1) translation equivalents (Kroll & Linck 2007; Kroll & Stewart 1994). After all, you may have only had 10 minutes of experience with the word *écureuil,* but you have had many years of experience with the word *squirrel*. This is to say that newly learned words do not have direct connections to their meanings at all, but instead indirectly access meaning via the L1 translation equivalent word, using the well-established connection between that word and its meaning. So the path you follow is *écureuil* → *squirrel* → SQUIRREL (where upper case indicates the meaning of *squirrel*). This is a specific instantiation of the idea of subordinate bilingualism (Weinreich 1953). Kroll and Tokowicz (2005) review the evidence for this view. For example, some experiments have compared the time to translate a word from L1 to L2 with the time to name a picture in L2. If translation were performed via meaning then these two tasks should be performed in about the same time – in both cases the appropriate meaning would be activated (by either an L1 word or a picture) and the appropriate L2 word retrieved. This is indeed the case at relatively high levels of proficiency (Potter, So, Von Eckardt & Feldman 1984). But at lower levels of proficiency L2 speakers are faster to translate from the L1 to the L2 than to name a picture in the L2 (Chen & Leung 1989; Kroll & Curley 1988). This could be explained by assuming that when translating they use a direct connection between the L1 and L2 words; in other words, a lexical-level shortcut that bypasses meaning. Only as proficiency increases do learners translate by accessing meaning (see Sunderman, this volume, for evidence). Subsequent experiments suggested that the tendency to translate using direct lexical connections is particularly strong in the L1 to L2 direction. Assuming that the links from meanings to L2 words are particularly weak then lexical connections from L1 translations provide an alternative method for retrieving L2 words. These assumptions became enshrined in the so-called "Revised Hierarchical Model," or RHM (Kroll & Stewart 1994; see Sunderman, this volume, for a discussion of the RHM).

Alternatively, a direct connection between *écureuil* and its meaning might be forged straight away. Presumably, when learning words through simple L1 translation equivalents, all you can do is attach the L2 word to the same meaning that is linked to the L1 word. This would mean that you do access meaning directly from the word, but it is the same meaning that was attached to the L1 word. So the path you follow is *écureuil* → SQUIRREL. And indeed, a number of researchers provide evidence that even *newly* learned L2 words do indeed access meaning directly (Altarriba & Mathis 1997; Finkbeiner & Nicol 2003; Sunderman, this volume). For example, there are semantic effects on L2 to L1 translation tasks (de Groot & Poot 1997) even for newly learned words (Altarriba & Mathis), and there are even stronger semantic effects when the new words are learned with

pictures rather than with L1 translations (Comesaña, Perea, Pineiro & Fraga 2009; Finkbeiner & Nicol).

Clearly there is a disparity between the findings of these studies. There are a number of factors that one might appeal to as an explanation for this disparity. One is whether the test tasks involve comprehension or production (see Sunderman, this volume, for related discussion). Evidence for use of L2 to L1 translation links tends to come from tasks that require translations to be produced, evidence for direct access of meaning comes from tasks such as translation recognition that merely require comprehension (i.e., are these two words correct translations?). Even the RHM assumes that L2 words have both direct lexical connections to L1 words and direct (but weak) connections to meanings, so it is possible that different tasks encourage reliance on different pathways. It could also be argued that the tasks that are being used to address this issue are not ideal because they force the person to access meaning, both in single word translation or translation recognition. Obviously the person who has just learned that *écureuil* means *squirrel* in a sense "knows" what the word means and can use this information in a variety of tasks. But if we looked at what aspects of the word's meaning become active spontaneously and automatically then we might expose more directly how the word is wired into the mind. By "automatically" we mean rapidly and without conscious intention, and even when not directly required by the task (see Segalowitz, Lacroix & Job, this volume, for a discussion of automaticity in L2 learning). This is where the priming methodology is useful.

Priming studies and L2 word learning

Priming, and specifically in the present case, semantic priming is a particularly useful method for probing how words are wired into the mind. In this paradigm participants are required to make some kind of decision about words, for example to indicate whether they are words or nonsense words – known as the lexical decision task (e.g., *blemp* would be a nonword in an English lexical decision task). Each word or nonword is preceded by a brief presentation of another word for perhaps as little as a quarter of a second. This is the prime. But the participant is only required to respond to the second word, the target, indicating whether it is a word or not. The crucial manipulation is the semantic relatedness of the prime and target (hence only the trials on which the target is a word are of interest). Sometimes the prime and target are semantically related (e.g., *chair-table*) and sometimes they are unrelated (e.g. *plant-table*). A robust finding is that target lexical decision times are faster after related primes than unrelated primes (Meyer & Schvaneveldt 1971). What makes this phenomenon so interesting is that it occurs automatically; that is,

without the person having a conscious intention to find a meaningful connection between the words. After all, the participant is not required by the task to process the prime words at all. Of course it is possible that they might come to understand what the experiment is about and try to use the prime words to anticipate the target (e.g., on seeing the prime *black* they might expect to see the target *white*). Indeed, in an elegant series of experiments Neely (1977) showed that participants can indeed do this, but only if the interval between the prime and target is long enough. When the interval is short, about 250 milliseconds (ms), the priming effect appears to be automatic. Indeed, semantic priming effects can even be obtained when the prime word is presented subliminally, that is, in such a way that it is not consciously perceived at all (Klinger & Greenwald 1995; Marcel 1983).

A prevalent metaphor for the process underlying automatic semantic priming effects is that of "spreading activation" around a semantic network (Collins & Loftus 1975). In this view, words are represented by nodes in a lexical network, and are connected to corresponding conceptual nodes in a semantic network (note that in this framework the terms "semantic" and "conceptual" are used interchangeably). Words that are semantically related are connected together in the semantic network, and the distance between the nodes is a reflection of their degree of semantic relatedness. In a priming experiment, when a prime word is recognised its lexical node becomes active, activation is passed to the corresponding conceptual node in the semantic network, and spreads to related nodes, with the amount of activation proportional to the distance between them. Activated nodes in the semantic network then pass activation back down to their corresponding nodes in the lexical network. If one of these words should be presented as the target, its recognition would be facilitated because its representation is "primed." Under appropriate conditions (see below) the spread of activation from the prime is assumed to be automatic. Therefore, this metaphor suggests that priming can be used to study the way that a word is connected to other words in the lexical and semantic networks.

Naturally, semantic priming has been used extensively to study processing of L2 words. Of particular interest has been whether the effect can be obtained across languages. For example, whether for an English learner of French *dog* primes *chat* (cat, L1-L2 priming), and whether *chien* (dog) primes *cat* (L2-L1 priming). If such effects were obtained under conditions where conscious expectancy strategies have been neutralised, and the spread of activation from the prime can be assumed to be automatic, then it would suggest that L1 and L2 words are directly connected to the same underlying conceptual system. Over the years a large number of studies have examined this question (for a review, see Altarriba & Basnight-Brown 2007). There is good evidence for automatic semantic priming from L1 to L2 in people of high L2 proficiency, for example,

when primes are presented very briefly (Keatley, Spinks & de Gelder 1994; Tzelgov & Eben-ezra 1992), when the task encourages them to ignore the primes (Fox 1996), or when the primes are presented subliminally; that is, so briefly that they cannot even be consciously perceived (Basnight-Brown & Altarriba 2007; Perea, Dunabeitia & Carreiras 2008). However, in many of these studies it was also found that the priming effect was much weaker (or not significant at all) in the L2-L1 direction (Basnight-Brown & Altarriba; Fox; Keatley et al.; Tzelgov & Eben-ezra). Only in one study of highly proficient bilinguals were the semantic priming effects the same size regardless of direction, and regardless of whether primes and targets were in the same or different languages (Perea et al.). Thus, whilst these studies certainly suggest that L1 and L2 words activate the same underlying semantic system, it appears that only at the very highest levels of proficiency are the semantic priming effects from L1 and L2 words equivalent.

Many of the studies that failed to find L2-L1 semantic priming did in fact find L2-L1 translation priming; for example, when *chien* primes *dog* (Basnight-Brown & Altarriba 2007; Fox 1996; Keatley et al. 1994; and see Altarriba & Knickerbocker, this volume, for L2-L1 translation priming from newly learned words). The contrast between L2-L1 semantic and translation priming could be taken as support for the RHM – L2 words are primarily connected to their L1 translations at the lexical level and do not have strong direct connections to meaning. However, two problems with this explanation are that weak L2-L1 semantic priming has been found even in bilinguals of high proficiency, at which point even the RHM assumes that direct connections from L2 words to meanings have been developed. Secondly, under the specific situation where primes are presented so briefly that they are not consciously perceived even L2-L1 translation priming seems to be obliterated, even though it persists in the L1-L2 direction (Jiang & Forster 2001). Thus, accessing meaning from L2 words might be less automatic than from L1 words even at high levels of proficiency.

What meaning is associated with L2 words?

An alternative is not that the strength of the connection from L2 words to meanings is weak, but simply that in some sense "less" meaning is associated with L2 words than L1 words. After all, L2 words have probably been encountered less often, and in a more restricted range of contexts, and this will presumably reflect the richness of the meanings that are associated with them. For example, Finkbeiner and Nicol (2003) argue that L2 words tend to be associated with a more restricted range of senses than L1 words, and that this could explain why under certain conditions even translation priming effects are weaker from L2-L1 than the reverse.

It may also be that there are restrictions on the "type" of semantic information that is associated with L2 words, particularly at lower levels of proficiency. In the above studies, as in most studies of semantic priming, the aim is simply to see whether prime words activate semantic information, where the term "semantic" is very vaguely construed as being concerned with meaning. In order to create sets of semantically related target words researchers typically turn to word association norms which list the words that most commonly come to mind when a person reads a single word (e.g., a common response to *knife* might be *fork*). But such responses can reflect a wide variety of different types of meaningful relationships. Of course one finds the standard semantic relations such as synonymy, antonymy, and hyponymy. But there are also pairs that are not highly semantically similar at all, such as *thunder-lightning*. As things in themselves, thunder and lightning are very different, in the sense that they do not have many shared properties, yet they are clearly strongly related in our minds. Not only do the words often occur together (they are frequent collocates) but so too do the corresponding things in the world. As another example, *squirrel-animal* is a classic lexical-semantic relation that depends on the intrinsic properties of SQUIRREL (its denotation), whereas *squirrel-nut* reflects an association between two concepts that derives from our general knowledge of squirrels. Furthermore, *squirrel-animal* is a relation that holds regardless of the context, whereas whether we think of a nut in relation to a squirrel might depend upon the context. Barsalou (1982) distinguished context-independent and context-dependent properties of concepts. For example, "can be walked on" is a property of a roof, but it may only occur to us in a specific context (of repairing one). On the other hand, "contains money" is a property of banks that we activate whatever the context, presumably because it is germane to our understanding of what a bank is. So even if we think of all of these properties and associations as being represented in the same semantic network, some will be more accessible than others depending on the context.

There is in fact some evidence that semantic similarity and association behave differently, at least when paradigms that measure highly automatic priming are used. Williams (1994) examined subliminal priming effects from L1 English to L2 French comparing semantically similar but not translation equivalent pairs, such as *hat-casque* (helmet), and associated but not semantically similar pairs such as *shoe-pied* (foot). A priming effect was obtained for the semantically similar pairs, but not the associates, even though the associates did produce robust priming when the primes were clearly visible. Other studies have also failed to find automatic cross-language priming between associates (de Groot & Nas 1991; Keatley & de Gelder 1992). However, because these studies also found priming between translation pairs it could be argued that the failure to find an effect for associates was due to a failure to access semantic information, with translations being

primed via lexical level links, as predicted by the RHM. This does not apply to the Williams study, though, because the priming effect for semantically similar pairs was clearly semantics-based. Instead, the results suggest a restriction on the type of semantic information that is shared between languages, preventing priming of associates.

Thus, when using semantic priming to explore the semantic information automatically activated by words we need to be specific about what kind of semantic information we are targeting. This might be particularly relevant when we are looking at priming from newly learned words, where the person's experience with those words will be obviously very limited. With this in mind, the present studies specifically examined priming from newly learned words to "associates," defined as pairs that are not highly semantically similar, but are related by virtue of actions and events in the world. These were compared with pairs of translation equivalence and semantic similarity in an effort to uncover the kinds of semantic information, if any, that are automatically activated by newly learned words.

The present study

In the first two experiments, participants learned novel French words by pairing them with their English translation equivalents. However, because our focus in this study is on the access of semantic information, we did not test for priming between a newly learned word and the translation equivalent with which it had been paired during learning. If a person had simply learned *écureuil* by pairing it with *squirrel*, then any subsequent priming from *écureuil* to *squirrel* could simply reflect an episodic memory for the learning experience, rather than any kind of lexical connection between the words. The notion of "episodic memory" will become important later. We will be using the term to refer to a memory system that stores personal experiences, as distinct from a "semantic memory" system that stores facts about the world and meanings of words.[1]

Suppose however that English is not the person's L1, but their L2. For example, the person's L1 might be Cantonese Chinese, their L2 English, and they might be learning French as a third language (L3). This is precisely the situation with many Cantonese learners of French and the way they are taught French in Hong Kong

1. The notion of "episodic memory" that we use here should also be distinguished from "autobiographical memory" – memory for personal experience that is linked to time and place (Tulving 2001). There are reasons for treating autobiographical memory as distinct from a more basic episodic memory system that records conscious experience. Our more restricted notion of episodic memory is also linked to the notion of the "episodic buffer" and episodic long-term memory (Conway 2001).

(through instruction in L2). After learning that *écureuil* means *squirrel* (an L3-L2 translation pair) we might then test whether *écureuil* primes 松鼠 (squirrel). If there is an effect, then it cannot be because of episodic memory (assuming that the person did not covertly translate into Cantonese during the learning phase). Rather it is more likely that it would reflect access to the concept SQUIRREL from *écureuil*. In other words, priming between L3 and L1 translation equivalents becomes a test of priming based on semantic similarity as opposed to episodic memory. In the first series of experiments, participants learned novel L3 French words by pairing them with L2 English translations. They then performed a priming test in which the newly learned L3 French words appeared as primes (for just 250 ms) followed by L1 Chinese targets. In Experiment 1, we contrasted automatic priming from newly learned L3 French words to two types of L1 Chinese targets: translation equivalents, and associates of low semantic similarity (e.g., *écureuil* = "squirrel"– 果仁 = "nut"). If there were no priming of either kind of target, then this would support the RHM. According to this model, only priming from *écureuil* to *squirrel* would be expected via the translation link formed between them during training. But suppose the newly learned French words become immediately linked to the same semantic representation as the L2 words with which they were paired in training – that is, *écureuil* links to SQUIRREL. According this view, the new L3 word simply inherits all of the semantic information associated with the L2 translation equivalent. Anything that is primed by *squirrel* should also be primed by *écureuil*. So *écureuil* should not only prime 松鼠 (squirrel), but also associates of *squirrel*, such as 果仁 (nut). On the other hand, if there is some limitation on the automatic semantic priming from newly learned words, as suggested by the Williams (1994) study, then while priming from *écureuil* to 松鼠 (squirrel) may be obtained, priming from *écureuil* to 果仁 (nut) might not.

The report of these studies is in two sections. The first section describes experiments in which participants learned L3 French words by seeing them paired with L2 English translations. The second section describes experiments in which L3 French words were learned in the context of pictures.

Learning L3 words with L2 translations

The first set of experiments comprised two main vocabulary learning experiments involving learning L3 (French) words and being tested through priming from the new learned L3 (French) words to L1 (Chinese) targets. There were also subsidiary control experiments to check for priming between the items using L2 (English) and L1 (Chinese) words. We summarise the common features of the design of the experiments here (see Cheung & Williams, in preparation, for full details).

Participants

In the vocabulary learning experiments, the participants were Cantonese (L1) university students with fluent English L2 who were following beginner-intermediate French classes in Hong Kong. Their mean age was 22.3 (19–29). There were male and female participants (ratio approximately 1:3). They had begun their study of English officially in the first year of primary school at approximately the age of 6, as is customary in the school system in Hong Kong. They had received English-medium instruction at university in Hong Kong. All participants had been taking French classes at their university or Alliance Française for at least one year.

The subsidiary control experiments for Experiment 1 and 2 were carried out in Cambridge (United Kingdom) and Hong Kong, respectively. The participants in the control experiments were university students with Cantonese L1 and fluent English L2. In Experiment 1, they were graduate students at the University of Cambridge. At the time of testing, they had spent at least one year living in the United Kingdom. The participants' mean age was 26.4 years (23–32). In Experiment 2, the participants were university students in Hong Kong. Their mean age was 25.4 years (18–31). In both subsidiary control experiments, the male to female ratio of the participants was approximately 1:1. In all experiments, the participants were studying a variety of subjects.

Vocabulary learning procedure

The participants were asked to study 22 French words (Experiment 1, and 20 in subsequent experiments) that we assumed would be new to them based on reference to French learning materials. They learned the French words using flashcards with a French word on the front and the corresponding English translation on the back. They worked strictly in the French-English direction. They studied all of the words for however many cycles were necessary to achieve 100% accuracy (this was an average of 6 cycles, taking about 20 minutes).

Priming test procedure

The test phase consisted of a lexical decision task (see Altarriba & Knickerbocker; Trofimovich & John, this volume, for additional studies using this task). On each trial, a French word was presented for 250 ms, followed by a Chinese word or nonword. The participants had to indicate whether the target was a Chinese word or not (lexical decision task). Half of the French prime words were those that had just been learned (we refer to these as "new" words). The other half were French

words that we assumed would be known (referred to as "known" words). The latter were included to see whether semantic priming effects could be obtained at least for French words already known by the participants. Pencil and paper tasks were used to establish that the participants really did not know the "new" words prior to the experiment, and that they already knew the "known" words. Data from inappropriately known or not known items were excluded on an individual by individual basis (accounting for an average of only 2.9% of the data).

Experiment 1

In the first experiment, we examined priming from L3 (French) primes to L1 (Chinese) targets. The critical trials were those in which the target word was a Chinese word. Two types of prime-target relationships were compared – translation equivalents and "associates" (words that are not highly semantically similar but related by virtue of the structure of actions and events in the world). There were also two types of prime: new French words that had just been studied in the experiment, and known French words. There were 22 primes of each type divided into two groups of 11 each (the relevance of the groupings will be explained shortly). Table 1 gives some examples (see Cheung & Williams, in preparation, for full details).

Table 1 only lists related pairs. Of course, to assess priming effects we need to compare the response time to trials on which the prime and target are related with those on which they are not. Unrelated pairings were formed by recombining half of the prime-target pairs within each condition. Equal numbers of participants were tested on each presentation list so that across participants each target appeared equally often in the related and unrelated conditions.

Table 1. Example items from Experiment 1

Prime	**Target**	
	Translation	**Associate**
Known		
femme/woman	女人 (woman)	裙子 (skirt)
étudiant/student	學生 (student)	課室 (classroom)
enfant/child	孩子 (child)	玩具 (toy)
New		
écureuil/squirrel	松鼠 (squirrel)	果仁 (nut)
tablier/apron	圍裙 (apron)	廚師 (chef)
menotte/handcuff	手扣 (handcuff)	警察 (police)

Table 2. Priming effects in Experiment 1 (L3-L1 priming)

	Translation		Associate	
	Known	New	Known	New
Priming (ms)	64*	60*	−8	−16

* $p < 0.05$

Turning to the results, for simplicity and brevity we shall concentrate on target response times in the lexical decision task (there were very few errors, the overall mean error rate being 4.2%). The critical measure of priming is the difference between target response times in the related and unrelated conditions, calculated by subtracting the mean response time in the related condition from the mean response time in the unrelated condition. The priming effects in each condition are shown in Table 2, along with the statistical significance of the effects, as calculated by paired-samples *t*-tests.[2]

The results are clear. Whilst there was priming of targets that are L1 translations of the primes, there was no priming of targets that were associates, both when the primes were French words already known by the participants, and when the primes were French words that had only just been learned. Indeed, it is remarkable that French words that had only just been learned produced the same size priming effect on L1 Chinese translation targets as French words that were already familiar.

With regard to the translation priming effect, our initial assumption was that since the targets were in Chinese, which was not the language the French words were paired with during training, then any priming effect would reflect semantic activation, rather than translation links in the lexical system, or simple episodic memories. However, if this was the case, why was there no effect for the associates?

An obvious possibility is that the associated targets were not sufficiently related to support a priming effect. After all, no effect was obtained even for the known words; so perhaps our method of choosing associated targets by intuition was not successful. To test for this, a subsidiary control experiment was run on 14 Chinese native speakers with an advanced level of English. They did the same lexical decision task as the participants in the main experiment, but this time the primes were translated into English. Since there was no vocabulary learning part to this experiment, there is no real distinction between "known" and "new"

2. An analysis of variance (ANOVA) showed that there was a significant interaction between relationship type and relatedness, both in an analysis by subjects $F_1(1, 18) = 7.50$, $p < .05$, and an analysis by items, $F_2(1, 80) = 7.34$, $p < .01$.

Table 3. Priming effects in Experiment 1, subsidiary control (L2-L1 priming)

	Translation		Associate	
	Known	**New**	**Known**	**New**
Priming (ms)	162*	151*	57	97*

* $p < 0.05$

words – all words were known (as verified by a post-test in which the participants were required to provide Chinese translations to all the English words). However, the naming of the conditions is preserved for comparison with the main experiment. The priming effects are shown in Table 3. There was significant priming in all conditions. Therefore, the failure to find L3-L1 priming for the associates in the main experiment cannot be because the associative relationships were not, in principle, strong enough to support an effect.

The contrasting results of the L3-L1 and L2-L1 priming experiments is important because it demonstrates that newly learned foreign language words do not simply inherit the semantic representation of the translation equivalents with which they were learned. Since the control experiment shows that, for example, *squirrel* primes 果仁 (nut), then if on learning that *écureuil* means *squirrel*, one would expect that if *écureuil* is simply wired to the same semantic representation as *squirrel*, then whatever *squirrel* primes, *écureuil* should prime as well. This is clearly not the case.

How, then, can the difference in behaviour of translations and associates in the main experiment be explained? One possibility is that the novel French words (and indeed the known French words) were not directly connected to semantics at all, as predicted by the RHM for low-level learners. The priming of Chinese translation targets would have to be assumed to reflect activation passing from L3 to L2 and then on to L1 by a chain of lexical level translation connections, bypassing the semantic system entirely.

An alternative hypothesis is that whilst the associates were sufficiently related to produce automatic priming in principle, as shown by the L2-L1 control experiment, in practice they might not have been sufficiently related to produce an effect at this low level of proficiency. In the next experiment, we therefore compared associates with a potentially stronger type of semantic relationship – words that are highly semantically similar, or nearly synonymous. If priming from newly learned words can only be supported by lexical level translation connections, then no effect should be obtained even for pairs that are of high semantic similarity so long as they are not potential translation equivalents. But if it is the nature of the semantic relationship that is important, then priming might be obtained for these more strongly related pairs.

Experiment 2

In Experiment 2, we compared automatic priming from newly learned French words to semantically similar and associated targets, where the two types of relation were defined as described above. Ideally, we would have used the same prime words as in Experiment 1 but it turned out that this was not possible because of the problems of finding semantically similar targets for all of primes. We therefore developed a new set of items, based on 20 French words for the "new" and "known" conditions, respectively. For each prime, a semantically similar and associated Chinese word was selected. Some superordinate relations were also included in the semantically similar condition (e.g., *écureuil–animal*), see Table 4 for examples. The results of the priming task are shown in Table 5.

The first result of note is that there were significant priming effects for semantically similar targets from both newly learned and known French word primes. This tells us that it is possible to obtain semantic priming effects even from newly learned words. Even though they had only just been learned through pairings with L2 (English) translation equivalents, these new French words were able to automatically activate a semantic representation and prime a semantically similar Chinese target. Note that the numerical size of the priming effect for semantically similar targets was similar to that obtained for translation targets

Table 4. Example items from Experiment 2

Prime	**Target**	
	Semantically similar	**Associate**
Known		
chaise/chair	椸化 (sofa)	書桌 (desk)
livre/book	書本 (dictionary)	作者 (author)
faim/hunger	胃口 (appetite)	餐廳 (restaurant)
New		
écureuil/squirrel	動物 (animal)	果仁 (nut)
papillon/butterfly	昆蟲 (insect)	花朵 (flower)
casque/helmet	帽子 (hat)	意外 (accident)

Table 5. Priming effects in Experiment 2

	Semantically similar		**Associate**	
	Known	**New**	**Known**	**New**
Priming (ms)	70*	60*	68*	9

* $p < 0.05$

Table 6. Priming effects in Experiment 2, subsidiary control (L2-L1)

	Semantically similar		Associate	
	Known	New	Known	New
Priming (ms)	136*	68*	48*	58*

* $p < 0.05$

in Experiment 1 (Table 2). This makes it plausible to argue that both effects were semantically based, and that there was no contribution of translation connections. After all, participants had learned the French words paired with English, rather than Chinese, words. Translation equivalence can be regarded as a case of very high semantic similarity.

When we turn to associate targets the picture is rather different. This time the associates of known words did produce a significant priming effect, but the associates of new words did not. Of course, the immediate question is whether for the new French words the associates were sufficiently strongly related to the primes to produce an effect. So, as in Experiment 1, a subsidiary control experiment was run on Chinese speakers of English to test priming from the English translations of the French words to the Chinese targets. The results are shown in Table 6. Note that the known/new distinction merely refers to the different sets of items. All primes were known by these participants. The results show that even the associates of the "new" primes support a significant priming effect.

Having ruled out lack of a sufficiently strong associative relationship as an explanation of the lack of priming for associates of newly learned French words we are faced with something of a paradox. Priming from newly learned L3 French words to semantically similar L1 Chinese targets clearly suggests that the newly learned French words rapidly and automatically contacted a semantic representation. But if this was simply the semantic representation of the L2 English translation equivalent, then the L3 French words should have primed everything that the L2 English words did. The fact that they did not suggests that L3 French words are not simply wired to the same semantic representation as the L2 English words with which they were learned.

The hypothesis we pursued in the remaining experiments was that the associates of words need to be learned through experience. That is, for the word *écureuil* to activate the concept NUT, it has to have become associated with this concept through experience. In this view, the fact that the subsidiary control experiments showed significant priming from L2 English to associates in L1 Chinese is simply because these English words, being more familiar, and experienced in a wider range of contexts, had been personally experienced with the relevant concepts.

In the second series of experiments, we used a different method for teaching the vocabulary to the participants. This time the new French words were paired with pictures (see Altarriba & Knickerbocker, this volume, who examined the learning of new L2 words by pairing them with pictures or L1 translation equivalents). Some of the time the picture was simply the appropriate object; for example, a flash card with the word *écureuil* on one side and a picture of a squirrel on the other ("single picture" cards). However, some of the time the picture also contained the associate that would appear in the priming task; for example, a picture of a squirrel holding a nut ("context picture" cards). Even though the presence of the associate (e.g., the nut) is entirely incidental to the task of learning the meaning of the word, will *écureuil* subsequently prime 果仁 (nut)? If it does, then it might suggest that associates need to be learned through experience, rather than inherited from their translation equivalents. The failure to find priming of associates in Experiments 1 and 2 could then be because the relevant information was not present as part of the learning experience of the participants. This would point to a broadly episodic view of word learning that stresses the information that was activated at the time the word was learned, rather than integration into an existing lexical-semantic system.

Learning L3 words with pictures

In this section, we describe two experiments in which L3 French words were learned by being paired with pictures (see Altarriba & Knickerbocker, this volume, for a discussion of the use of pictures in vocabulary learning). For the first two cycles of training, the picture was simply the appropriate object; for example, a flash card with *écureuil* on one side and a picture of a squirrel on the other. This ensured that the participants knew which object was referred to by the word. In the remaining cycles, the pictures also contained the associate that would appear in the priming task (e.g., a picture of a squirrel holding a nut). As in Experiments 1 and 2, the participants for Experiments 3 and 4 were Cantonese (L1) university students with fluent English L2 who were following beginner-intermediate French classes at the University of Hong Kong. Their mean age was 22 years (19–27) and the male to female ratio was approximately 1:2. All participants had been taking French classes at the university for at least one year. The participants were studying a variety of subjects.

Experiment 3

The experimental material consisted of 40 prime-target word pairs that were semantically similar and associatively related. All stimuli, except for five prime-target pairs, were identical to the stimuli in Experiment 2. The five pairs were replaced either due to the prime being an abstract word or a word that was difficult to make clear in a picture. For each of the 20 words in the "new" condition, three flash cards were created with the French word on one side and a picture on the other. On the "single picture" cards, the picture only depicted the word. For the two "context picture" cards, the associate items were also included in the photograph or picture (e.g., each card showed a different depiction of a squirrel holding a nut, with a different squirrel and nut in each picture). All photographs and pictures were obtained online from photograph and image websites. The lexical decision task was designed in the same way as in Experiment 2.

In the study phase, the participants were given three sets of 20 flashcards. They were told to learn the 20 words using the pictures printed on the back of each card. The participants went through the set of "single picture" cards twice before studying the two sets of "context picture" cards. The procedures of the test phase (lexical decision task) were identical to Experiment 2. The reaction time data from the lexical decision task were analysed in the same way as in the previous experiments. The priming effects in each condition are shown in Table 7 below.

Consistent with Experiment 2, the semantically similar targets showed significant priming. With regard to the associates, we found, for the first time, a significant priming effect when the primes were newly learned French words ("new" condition). However, there was no effect for previously known French words.

With regard to the priming of associated targets from newly learned words, we would like to argue that this is because during the study phase the associated object was pictured along with the picture of the French word. The presence of this object was actually irrelevant to the task of learning the meaning of the French words, but nevertheless it became incidentally associated with it. The lack of priming of associates of "known" words would presumably be because these beginner learners of French have not experienced the words in the appropriate contexts for the relevant associations to be formed, even though they claimed to

Table 7. Priming effects (by subject) in Experiment 3

	Semantically similar		Associate	
	Known	New	Known	New
Priming (ms)	87*	46*	0	74*

* $p < 0.05$

know what the words mean in the post-test. Note that these "known" words are with a few exceptions the same as those used in Experiment 2, where significant priming of associates was obtained. But different people, with different personal experiences, participated in the two experiments, so we assume that it is the different usage history of the words for the two groups that accounts for the different results.

An immediate objection to our interpretation of the results might be that priming from newly learned words to associates emerged in this experiment simply because the words were learned with pictures, and not because some of those pictures depicted the relevant associates. Perhaps the pictures of the objects themselves encouraged deeper processing of the meanings of the novel French words, thereby making it more likely that associated concepts became linked to it. For example, in a study on 10 and 11 year-old children, Comesaña et al. (2009) found stronger evidence that newly learned words activated semantic representations when those words had been learned with pictures, as opposed to L1 translation equivalents (but see Lotto & de Groot 1998, for no such difference in adults, and Altarriba & Knockerbocker, this volume, for evidence that translation priming effects are just as large after words are learned with translations as after having been learned with pictures). Note, however, that if studying words with pictures simply forged stronger word-to-concept links, then we would have also expected to find stronger priming of semantically similar targets in this experiment compared to Experiment 2. A comparison of Tables 5 and 7 shows that this was not the case. The priming effects were numerically very similar. However, this comparison is only suggestive because different items were used in the two experiments.

Another potential objection to our explanation of the results might be that by including the associated objects in the study pictures we are in fact strengthening the association between the relevant concepts in semantic memory, rather than forming an association between the novel word and the associated concept. For example, seeing pictures of squirrels holding nuts strengthens the association between the SQUIRREL and NUT conceptual nodes in the semantic network, if only over the short timescale of the experiment. If this were the case, the newly learned word *écureuil* would be priming the associated targets by capitalising on a strengthened link in the semantic network. In contrast, our episodic view of word learning claims that associations are unique to specific words – it is the fact that the word *écureuil* has been experienced in the same context as a squirrel holding a nut that enables *écureuil* to subsequently prime 果仁 (nut). In Experiment 4 we tested this specific proposal.

Experiment 4

In this final experiment, participants again learned the novel French words with pictures. As in Experiment 3, two out of the three cards for each word contained an object that would appear as the "associated" target in the subsequent priming task. However, this time the object was not already a strong associate of the word. For example, a squirrel was depicted amongst some leaves on the ground, or a butterfly was depicted resting on a person's finger. Our question was whether, in the subsequent priming task, *écureuil* (squirrel) would prime 樹葉 (leaves), and *papillon* (butterfly) would prime 手指 (finger). Such an effect could only derive from the incidental pairing of the relevant objects in the study phase, and would show that priming from newly learned words to associates can reflect the specific study experience. It would also suggest that priming of associates of newly learned words in Experiment 3 was not simply because of overall deeper semantic processing from having learned the French words with pictures.

In addition, in the priming task we now not only test for priming between L3 *écureuil* and L1 樹葉 (leaves), but also between L1 松鼠 (squirrel) and L1 樹葉 (leaves). If seeing some pictures of squirrels with leaves simply forms a novel (or strengthens an existing) conceptual link between SQUIRREL and LEAF in semantic memory, then priming should be obtained regardless of whether the prime is in French or Chinese. But if what is formed is a specific association between the word *écureuil* and the concept LEAF then priming should only be obtained when *écureuil* is used as the prime; that is, there should be priming from *écureuil* to 樹葉 (leaves) but not 松鼠 (squirrel) to 樹葉 (leaves). Of course, we can only make this comparison because novel associations are used in this experiment. If we had used our staple *écureuil–nut* association, then we would have obviously expected priming from 松鼠 (squirrel) to 果仁 (nut) simply by virtue of the pre-experimental association between them.

The experimental material consisted of 40 prime-target word pairs that were semantically similar and associatively related. All prime words and their semantically similar targets were the same as in Experiment 3, but new associates were selected following the principles discussed above. The procedure in the study phase was the same as in Experiment 3. In the test phase, half of the subjects began with the L3-L1 version of the lexical decision task followed by the L1-L1 version. The other half of subjects undertook took the task with reverse language order. The results are shown in Table 8.

As expected on the basis of previous experiments, the semantically similar targets were primed in all conditions. Indeed, in the "new" condition, the French words appear to produce a numerically larger effect than the Chinese words

Table 8. Priming effects (ms) (by subject) in Experiment 4

	Semantically similar		Associate	
	Known	New	Known	New
French–Chinese	61*	81*	105*	54*
Chinese–Chinese	43*	42*	65*	10

* $p < 0.05$

(81 ms vs. 42 ms, respectively), although the interaction between these effects was not statistically significant. Nevertheless, after an average of just 6 cycles of training with flash cards, novel French words produce priming effects for semantically similar Chinese targets that are at least as strong as for the equivalent Chinese words.

With regard to the new associates it can be seen that, as in Experiment 3, priming from newly learned French words was obtained. Because these associations were themselves novel, we know that the effect must have derived from the specific experience of seeing the relevant objects paired in the study phase, even though these pairings were entirely incidental to the task of learning the meanings of the French words. Given this result, it seems likely that the priming of new associates in Experiment 3 was also because the relevant objects had been paired in the study phase. Priming of associates therefore seems to be a reflection of the study experience, rather than previous connections in semantic memory.

But what kinds of connections are being formed? Are they connections between concepts (e.g., between the concepts SQUIRREL and LEAVES)? Or between the new French word *écureuil* and LEAVES? To answer this question, we need to look at the Chinese–Chinese condition, where in contrast to the French–Chinese condition, there was no significant priming effect.[3] So the priming effects are language-specific, and presumably reflect the fact that it was the French word *écureuil* that was learned in the context of a squirrel amongst some leaves. This novel association was highly language-specific.

Finally, the known associates produced significant priming regardless of whether the primes were in French or Chinese. The result for the French primes contrasts markedly with that obtained in Experiment 3 where exactly the same materials were used except for one pair. Clearly, the effects in this condition are

3. An ANOVA showed that for the New Associate condition the interaction between relatedness and prime language (French or Chinese) was significant in an analysis by subjects, $F_1(1, 22) = 4.41$, $p < .05$, and approached significance in an analysis by items, $F_2(1, 18) = 3.57$, $p = .075$.

extremely unstable across experiments (compare also Experiments 1 and 2). We shall return to this point below.

Discussion

The aim of this study was to examine the nature of the meaning that is automatically activated by newly learned words. On initially learning an L2 word as a translation equivalent of an L1 word, we may feel that we know its meaning, but is this the same meaning that is associated with the L1 word? What kind of meaning does the L2 word immediately and automatically activate when we read it, and how does that differ from the meaning activated by the L1 equivalent? We investigated these issues by using the semantic priming paradigm. Our participants, native speakers of Cantonese Chinese, learned novel L3 French words as translations of L2 English words and were subsequently tested on semantic priming from the L3 French words to L1 Chinese targets. In Experiment 1, there was priming from newly learned L3 French words to L1 Chinese translation equivalents; and in Experiments 2 to 4, newly learned L3 French words primed semantically similar L1 Chinese words. Clearly, newly learned words are able to produce semantic priming effects. But in terms of the frameworks for thinking about the bilingual lexicon that were reviewed earlier this finding in itself is not very indicative. On the RHM, the L3 word could be connected directly to its L2 translation equivalent at the lexical level, and access the meaning associated with the L2 word from there. Or else the L3 word could be directly connected to the meaning of the L2 word. The important point is that both of these approaches imply that at these initial stages of learning, the L3 word essentially activates the same meaning that is associated with the L2 translation equivalent (Jiang 2000, 2002). So whatever the L2 word primes, the L3 word should prime as well, assuming that the L3 words activate any meaning at all. However, our experiments show that this is not the case. Experiments 1 and 2 also tested for priming of "associates" of the L3 words – concepts that are not in themselves semantically similar, but which are related by virtue of the structure of events in the world. In Experiments 1 and 2 there was no priming from the newly learned L3 French words to associates presented in L1 Chinese, even though control experiments showed that the same associative relations supported priming from L2 English to L1 Chinese. We conclude that whilst newly learned words do rapidly access meaning, they do not simply inherit all of the semantic information associated with the translations with which they were paired during learning. This would appear to contradict the prevalent assumption that newly learned words access the same meaning as their translation equivalents (Jiang 2000; Kroll & Stewart 1994).

The hypothesis we pursued in the remaining experiments was that associative relations of this type need to be learned by actually experiencing the words and relevant concepts, and not by inheritance from the translation equivalent. Experiment 3 supported this idea by showing that when a novel L3 French word is learned by being paired with a picture, and when the picture also contains an associated object (e.g., a squirrel holding a nut), then that French word will subsequently prime recognition of the L1 Chinese label for the associated object. Experiment 4 showed that this effect does not arise simply because the association between the objects is strengthened within the semantic system. This time the pictures depicted the relevant object in the context of an object with which it is not already strongly associated (e.g., a squirrel amongst leaves). In the subsequent priming task, there was L3 French to L1 Chinese priming for this novel association, but no L1 to L1 priming. This demonstrates that the association is specific to the L3 French word.

We believe that these results point to an essentially episodic view of word learning in which the emphasis is not upon integration into the existing linguistic system, but upon establishing language-specific associations between form and meaning through experience. This seems like an obvious, and simple, view of word learning, and we believe that it is. But it does seem to challenge prevalent views of the bilingual lexicon that stress linkages to pre-existing L1 representations, be they L1 lexical representations (translation connections) or L1 semantic representations (direct mapping to meaning). Instead, our view stresses the importance of individual learning episodes in providing the meanings with which words are associated.

The importance of episodic memory (in the sense defined earlier) in vocabulary learning has been alluded to previously in the literature. Jiang and Forster (2001) proposed that L2 word forms are represented in episodic memory, where they are linked to episodic memory representations of L1 word forms (i.e., the lexical level translation connections from L2 to L1 proposed by the RHM are conceived of as represented in episodic memory, as opposed to lexical memory). This proposal was based on the finding that whereas subliminally presented L2 words do not prime lexical decisions to L1 translation equivalents, they do prime L1 translation equivalents in an episodic recognition task (participants first studied a list of L1 words, and in the subliminal priming task had to indicate whether the target word was in the studied list or not). Assuming that the episodic task taps episodic memory and the lexical decision task taps lexical memory, then the implication is that even for these quite advanced bilinguals, L2 forms and their connection to L1 forms are represented in episodic memory. Perfetti, Wlotko and Hart (2005) propose that newly learned word forms (and possibly also meanings) are encoded in episodic memory based on the fact that

they produce event-related potential signatures (patterns of brainwave activity) that are characteristic of episodic recognition, even though the participants were performing a semantic judgement on the words. What both of these studies point to is the idea that processing of L2 words is somehow closely related to specific experiences (see also Altarriba & Knickerbocker, this volume, for a formulation of this idea in the context of transfer appropriate processing). However, we would question whether it is necessary to postulate an episodic memory system that is separate from the lexical and semantic systems.

As a framework for thinking about the kind of memory processes that we have in mind it is useful to consider "Multiple-Trace" (Hintzman 1986) or "Connectionist" (McClelland & Rumelhart 1985) models of memory. The spirit of these approaches to memory is similar, but here we will focus on Hintzman's multiple-trace memory model. According to this model, each experience forms a separate trace in memory. For example, a child might hear the word *doggie* at the same time as seeing a spaniel in the local park, and this experience is laid down as a unique memory trace. If the child were subsequently to hear the word *doggie*, this would activate the whole memory trace, and so an image of the spaniel in the park might come to mind. If the child subsequently heard *doggie* while seeing other dogs in other contexts (e.g., a labrador at home, a collie at auntie's house, etc.), then each of these would lay down a unique memory trace. On hearing the word *doggie*, the response from memory would be the average of all memory traces in which that word is encoded; something approaching a prototype representation of a dog, although entirely dependent upon the sample of specific dogs in this child's experience. Because memory traces are encoded using a common repertoire of features it is possible to average over them to construct prototypes at the moment of retrieval. Conversely, on seeing a terrier, this experience would resonate with memory traces involving other dogs to the extent to which it matches them, and the word *doggie* might come to mind more or less strongly depending on the degree of match (hence, *doggie* might be less likely to come to mind on seeing a chihuahua). Given the right cue, recall of specific information is also possible. On seeing a spaniel in the street, the specific memory trace of the spaniel in the park might come to mind. Thus, in this kind of system, memory is only composed of traces of experience, the response of memory to a cue is simply a function of the match between that cue and the contents of memory traces, and whether the "echo" from memory is relatively abstract (e.g., a prototype dog) or specific (e.g., the memory of a spaniel in the park) depends upon the specificity of the cue.

Hintzman (1986) showed how a model of memory built on these simple principles could simulate basic findings from semantic memory research, such as

prototypicality and category size effects.[4] The multiple-trace model also provides a simple explanation of our results. When a participant in Experiments 1 and 2 learned that *écureuil* means "squirrel," they formed memory traces containing the word *écureuil*, the word *squirrel* and the meaning activated by *squirrel*. By "meaning" here we mean the semantic features that are regarded by the participant as being relevant to the word learning task, which were presumably the features associated with the denotation of the word *squirrel*. In this sense, the novel word does become linked to the meaning of its translation equivalent, but the link is not to a representation of meaning in a "semantic network" (which does not exist in this view), but to whatever aspects of that meaning happen to enter into the trace of the learning episode. In the priming task, when *écureuil* was presented, it reactivated these memory traces, causing the semantic features that were encoded along with *écureuil* to become re-activated as well. These features would have primed the Chinese translation equivalent of *écureuil* because these same features are contained in memory traces for the Chinese word 松鼠 (squirrel). The same applies to priming of a semantically similar Chinese word like 動物 (animal). But there is no reason why 果仁 (nut) would have been primed since we assume that this concept would not have entered into the memory traces of the learning episodes. There is no reason why somebody learning that *écureuil* means squirrel should think of nuts.

In contrast, in Experiment 3, the features of NUT would have been represented as part of the memory traces of the learning episodes. Subsequent presentation of *écureuil* in the priming task re-activated these traces, thereby activating the concept NUT, and priming 果仁 (nut). This does not mean that the person thinks that the features of NUT are part of the meaning of *écureuil*. The concept NUT just happens to be present as part of the context in which the word was learned. Similarly, in Experiment 4, *écureuil* became associated with a novel concept, LEAVES, producing priming from *écureuil* to 樹葉 (leaves). But the Chinese word 松鼠 (squirrel) will not prime 樹葉 (leaves) because the memory traces activated by the prime word 松鼠 (squirrel) do not contain that concept, or at least not frequently enough to produce a priming effect.[5]

4. Clearly, though, as a model of word learning this approach is rather naïve. In particular one needs to take into account the important role of the child's hypotheses about what a novel word might mean, as determined by joint attention (Bloom 2000), or principles such as mutual exclusivity and the basic level constraint (Markman 1989). Such factors could be regarded as influencing the way in which the experience is encoded in memory.

5. Our approach is similar to the Distributed Feature model proposed by van Hell and de Groot (1998). Their model also stresses the uniqueness of the mapping from L1 and L2 words to a common set of semantic features. However, like other distributed models of semantic memory

The Multiple-Trace model is also consistent with the results of other bilingual priming experiments in the literature. It is not surprising that translation priming is robust across languages. If words have been learned as translation pairs, then there will be memory traces containing both the L1 and L2 word forms, equivalent to the lexical link in the RHM, or the links in episodic memory proposed by Jiang and Forster (2001). But we also propose that when a word is learned by pairing with an L1 translation, the memory trace also contains the semantic information activated by the L1 word, and as the L2 word is encountered in more contexts further traces containing aspects of meaning will be laid down. Thus, even though the specific traces of pairings of L2 and L1 forms laid down during learning might provide a basis for the notion of form-level translation connections (as in the RHM) it seems likely that translation priming effects, at least using clearly visible primes, are largely driven by the high level of overlap in meanings between L2 and L1 words. Although this will result in priming in both directions, it would also be expected that priming would be stronger from L1 primes than from L2 primes. This is simply because L2 words have been encountered in a more restricted range of contexts, and hence activate "less" meaning when projected to the memory system. Or, as Finkbeiner, Forster, Nicol, and Nakamura (2004) argue, L2 words are associated with a restricted range of senses, which reduces their degree of overlap with the meanings of L1 words (see also Segalowitz et al., this volume, for evidence that word processing is less "information rich" in L2 than L1). As a result, L2-L1 translation priming would generally be expected to be smaller than the reverse, which has been observed both for visible (Fox 1996; Keatley et al. 1994) and, in an extreme form, for masked primes (Jiang & Forster, using lexical decision). Nor should it be surprising that semantic priming is not nearly as robust across languages as translation priming. If, as we assume, semantic priming depends upon having experienced the prime word in a situation where the target concept was also active, then the probability of this for a particular set of associations and learners would be highly variable.[6] Again, priming will be more robust in the L1-L2 direction (Basnight-Brown & Altarriba 2007; Fox; Keatley et al.) simply because of the range of contexts in which L1 words have been encountered,

(Masson 1995), this model can only account for priming by virtue of semantic overlap. It does not distinguish between semantic features that are part of the denotation of a word and features that are part of its context of use. The Multiple-Trace model has the potential to deal with this problem because the memory trace can have internal structure that reflects the person's understanding of the event at the time of encoding.

6. Note that this depends upon the nature of the association. Many associates contain high degrees of semantic overlap (e.g., boy-girl) and so would be expected to be less dependent upon experiencing the word in a context where the associate is present.

increasing the chances that the particular association targeted in the experiment will be present in a subset of memory traces containing the L1 word. However, at very high levels of proficiency, this asymmetry should disappear (Kotz & Elston-Güttler 2004; Perea et al. 2008). In the present study, the "known" French words only produced L3-L1 priming of associates in two out of four experiments. Given that the participants were beginner learners of French this can be attributed to a lack of homogeneity in their experience of these words. Note, however, that priming effects were obtained for the same associative relations in our L2 (English) to L1 (Chinese) subsidiary experiments, which we attribute to these participants' greater experience with the English words.

As noted in the introduction, previous research has tended to find small, or in many cases, non-existent, semantic priming effects in the L2-L1 direction. Indeed, Silverberg and Samuel (2004) only found L2-L1 priming in participants who learned the L2 early in life, and not for late learners of an equally high level of proficiency. They go as far as to argue that only early learners develop a combined conceptual system for their two languages, while later learners have distinct conceptual systems. The problem with this view is that it cannot explain why L1-L2 semantic priming is readily obtained even in late learners (Williams 1994) and even in the absence of L2-L1 priming (Basnight-Brown & Altarriba 2007; Fox 1996; Keatley et al. 1994). If the conceptual systems are distinct in late learners then it should not be possible to obtain semantic priming in either direction. A more likely explanation of the asymmetry, and the one we argue for above, is that L2 words have been encountered in a more restricted range of contexts than L1 words, and hence do not activate the same range of semantic information.

Even so, given that previous studies have only obtained L2-L1 priming in more proficient learners, it may seem surprising that in our experiments newly learned words produced highly robust priming of semantically similar targets in Experiments 2–4, and of associated targets in Experiments 3 and 4. There are two possible explanations for this. The first is that experiments that train participants on a small set of novel words with concentrated exposure actually create conditions of very high proficiency, but in a highly restricted domain (see Friederici, Steinhauer & Pfeifer 2002, for the same argument applied to grammar). All of the relevant memory traces were recently formed; unlike in naturalistic acquisition, a very high proportion of them (two thirds) contained the relevant association, and all participants had the same experience with the words. We would argue that it was this homogeneity of experience that allowed us to observe an effect that may be more difficult to obtain "in the wild."

An alternative possibility is that semantic priming from L3 words (and in the control experiments, from L2 words) occurred because our priming procedure did not tap into the same level of automaticity of priming effects as in some other

studies. A problem in this area of research is that whilst all studies claim to measure priming effects that are automatic, in the sense of not being contaminated by conscious strategies, they vary widely in the methods employed to achieve this (Altarriba & Basnight-Brown 2007). Manipulations here include subliminal prime presentation, continuous list presentation, which obscures the pairings of primes and targets, and manipulations of proportion of related trials in the experiment. Here we used a prime-target interval that we assumed would be too short to permit conscious expectancies to develop (following Neely 1977), but the primes were visible, the pairings of primes and targets was perfectly evident, and the proportion of related trials was relatively high. It is possible, therefore, that our procedure did not require the same very high levels of automaticity of semantic access from the primes as procedures used elsewhere in the literature. It may well be therefore, that only words learned early in life, or of some high level of familiarity, can produce priming effects under conditions that require the highest levels of automaticity of lexical access. Our new French words were very unlikely to have been learned to this level of automaticity, but we would argue that this does not alter the force of our essential finding that priming of associated targets was critically dependent upon the learning experience of the individual. Our experiment tapped into the relationship between priming effects and experience, but did not induce the level of automaticity that would come from years of language experience.

Pedagogical implications

Our first two experiments show that after relatively little training on novel L3 French words paired with L2 English translations, the novel words produced automatic semantic priming of translations and semantically similar words in the L1 Chinese. This would suggest that the "paired associate" method of learning does not just encourage lexical-level connections between words, but also results in rapid access to meaning (see Altarriba & Knickerbocker and Barcroft, Sommers, & Sunderman, this volume, for examples of studies investigating a similar method of learning L2 words). Does this make it a good method for learning vocabulary?

Modern foreign language teaching practice tends to eschew the use of translations for learning vocabulary. The basis for this view is presumably that it is assumed that if words are learned by translation the learner will simply associate the L1 meaning with the L2 word, either directly, or via a translation connection; a view that we discussed in the Introduction and one that is shared by psycholinguistic models of word learning such as the RHM. Since it is clearly not the case

that all aspects of L1 word meanings are always shared by L2 translations, the fear is that L1 usage patterns will exert an undue influence over L2 comprehension and production.

However, our experiments suggest that in fact when words are learned by translation not all of the meaning of the L1 word is inherited. We have argued that only the meaning that is activated by the stimuli in the learning episode, be they pictures or words, become associated with the novel word. So when learning via translations, this will presumably be what the learner assumes to be the denotation of the L1 word, or its core meaning. The contextually-dependent aspects of meaning, such as associations with contexts of use and collocation patterns, would be predicted to be incidentally learned through exposure to the word in context by amassing very many more memory traces. The initial learning of the core meaning provides only the seed for this longer term learning process. Or to borrow a term from the statistical learning literature, the use of translation equivalents "bootstraps" the learning process and allows the L2-specific nuances of meaning to be acquired over time (see de Groot & van Hell 2005, for a similar view).

But this argument only applies to the context-dependent aspects of word meaning. The problem is, of course, that if the core meanings of the L1 and L2 words are not identical, then the learner is in danger of associating the wrong meanings with the L2 word, will activate inappropriate meanings in comprehension, and will have a "semantic accent" in production. And there is indeed good evidence for interference from L1 meanings during L2 processing even in advanced bilinguals (Elston-Güttler & Williams 2008; Jiang 2002; Thierry & Wu 2007). For example, Elston-Güttler and Williams showed that when German learners of English read a sentence like "On his foot was a large bubble," they find it harder to say that *bubble* is an unacceptable completion than do non-German speakers. This is because the "translation equivalent" of *bubble* in German is *Blase* which has the additional meaning BLISTER, which would make an acceptable sentence completion. Importantly, this effect was obtained in advanced learners of English who knew full well the meaning of *bubble*. In other words, they had acquired the correct form-meaning mapping but still showed interference from the L1 translation of the word during processing. The ubiquity of such effects in both comprehension and production (Grosjean 1982) might make one nervous about using translation equivalents during word learning.

Of course, this is to assume that the way a word is learned has a lasting impact on the way that it is processed. Is it really the case that learning a novel word as a translation equivalent forges such a strong translation link between them that it influences processing years later? Recall that the RHM posits strong L2-L1 lexical level connections between translations. But it is assumed that these are not just the

result of early learning experiences with translation pairs, but that they are built up over time because the learner finds it easier to exploit the stronger L1 form-meaning mappings in comprehending L2 words. So long as the L2 form-meaning mappings are relatively weak, translation, according to the model, provides the preferred route for accessing meanings from L2 words. Learning words as translations might rapidly establish a lexical level connection that then becomes reinforced through subsequent processing. Or learners may come to forge such connections spontaneously regardless of how they are taught.

Our episodic view is somewhat different and would appear to lay more emphasis upon learning experiences. Learning words as translation pairs will establish a lexical connection between them in memory, which may or may not influence subsequent processing in the way just described. This is a question for future research. However, and perhaps more importantly, it also establishes a memory trace of the association between the L2 word and the *meaning* of the L1 word *that was active at the time*. It is the nature of the meaning of the L1 word that was active at the time that becomes critical. If the German learner of English simply learns that *blister* means *Blase*, then there is no reason why, in that moment, they should not assume that the meaning of *blister* extends to both BLISTER and BUBBLE. Thus, the real danger of learning vocabulary through translation pairs lies not so much in establishing lexical level connections, but in encouraging false assumptions about the meanings of L2 words.

For this reason, it is clearly preferable to learn vocabulary with pictures and realia wherever possible simply because these are less ambiguous, and because they avoid the potential pitfall of forming lexical-level translation connections (even though, as noted above, we simply do not know to what extent instructional method is to blame for reliance on translation). However, this is obviously difficult for abstract words. This is unfortunate, because it is here that polysemy is rife and the lack of translation equivalence most common, making learning by translation perilous. One solution to this problem might be to ensure that when words are learned with L1 translations (e.g., *blister-Blase*) the L1 word is presented in a context that makes the relevant meaning apparent (e.g., *Er hatte eine grosse Blase auf seinem Fuss*, "He had a large blister on his foot"), and for this to be contrasted with an example where a different word would be used in the L2 (*Das Getränk war voller Blasen*, "The drink was full of bubbles"). Having established the relevant meaning, it would be important to rapidly move on to examples of the novel word in L2 contexts so as to avoid the formation of lexical level translation connections. Of course, this depends on a sophisticated semantic analysis of the relationship between L1 and L2 word meanings, and instruction that is tailored to specific L1s. For this reason, many would argue that the meanings of abstract words can be worked out from contexts that are expressed in the L2. We shall not enter into

the debate over the relative merits of guessing from context methods (de Groot & van Hell 2005; Pressley, Levin & McDaniel 1987; Sternberg 1987), although we would point out that any method should induce sufficient engagement with both the form and the meaning of the target word. If effort is to be expended anywhere, it should perhaps be at the level of form retention, the relationship between form and meaning, and exemplifying L2 usage patterns, rather than in working out meanings that the learner already knows. Thus, translation may be an efficient means of seeding the larger word learning process, but it needs to be used with great care.

References

Altarriba, J., & Basnight-Brown, D. M. (2007). Methodological considerations in performing semantic- and translation-priming experiments across languages. *Behavior Research Methods, 39*, 1–18.

Altarriba, J., & Mathis, K. M. (1997). Conceptual and lexical development in second language acquisition. *Journal of Memory and Language, 36*, 550–568.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory and Cognition, 10*, 82–93.

Basnight-Brown, D. M., & Altarriba, J. (2007). Differences in semantic and translation priming across languages: The role of language direction and language dominance. *Memory & Cognition, 35*, 953–965.

Bloom, P. (2000). *How Children Learn the Meanings of Words*. Cambridege, MA: The MIT Press.

Chen, H., & Leung, Y. (1989). Patterns of lexical processing in a nonnative language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 316–325.

Cheung, A., & Williams, J. N. (in preparation). Priming and early word learning.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407–428.

Comesaña, M., Perea, M., Pineiro, A., & Fraga, I. (2009). Vocabulary teaching strategies and conceptual representations of words in L2 in children: Evidence with novice learners. *Journal of Experimental Child Psychology, 104*, 22–33.

Conway, M. A. (2001). Sensory-perceptual episodic memory and its context: autobiographical memory. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 356*, 1375–1384.

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., & Poot, R. (1997). Word translation at three levels of proficiency in a second language: The ubiquitous involvement of conceptual memory. *Language Learning, 47*, 215–264.

de Groot, A. M. B., & van Hell, J. G. (2005). The learning of foreign language vocabulary. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of Bilingualism* (pp. 9–29). Oxford: Oxford University Press.

Elston-Guttler, K., & Williams, J. N. (2008). L1 polysemy affects L2 meaning interpretation: Evidence for L1 concepts active during L2 reading. *Second Language Research, 24*, 167–187.
Finkbeiner, M., & Nicol, J. (2003). Semantic category effects in second language word learning. *Applied Psycholinguistics, 24*, 369–383.
Finkbeiner, M., Forster, K., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language, 51*, 1–22.
Fox, E. (1996). Cross-language priming from ignored words: Evidence for a common representational system in bilinguals. *Journal of Memory and Language, 35*, 353–370.
Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial language processing: Evidence challenging the 'critical period' hypothesis. *Proceedings of the National Academy of Science, 99*, 529–534.
Grosjean, F. (1982). *Life with Two Languages*. Cambridge, MA: Harvard University Press.
Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93*, 411–428.
Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics, 21*, 47–77.
Jiang, N. (2002). Form-meaning mapping in vocabulary acquisition in a second language. *Studies in Second Language Acquisition, 24*, 617–637.
Jiang, N., & Forster, K. I. (2001). Cross-language priming asymetries in lexical decision and episodic recognition. *Journal of Memory and Language, 44*, 32–51.
Keatley, C. W., & de Gelder, B. (1992). The bilingual primed lexical decision task: Cross-language priming disappears with speeded responses. *European Journal of Cognitive Psychology, 4*, 273–292.
Keatley, C. W., Spinks, J. A., & de Gelder, B. (1994). Asymmetrical cross-language priming effects. *Memory & Cognition, 22*, 70–84.
Klinger, M. R., & Greenwald, A. G. (1995). Unconscious priming of association judgments. *Journal of Experimental Psychology: Learning Memory and Cognition, 21*, 569–581.
Kotz, S. A., & Elston-Güttler, K. (2004). The role of proficiency on processing categorical and associative information in the L2 as revealed by reaction times and event-related brain potentials. *Journal of Neurolinguistics, 17*, 215–235.
Kroll, J. F., & Curley, J. (1988). Lexical memory in novice bilinguals: The role of concepts in retrieving second language words. In M. Gruneberg, P. Morris & R. Sykes (Eds.), *Practical Aspects of Memory* (Vol. 2, pp. 389–395). New York, NY: Wiley.
Kroll, J. F., & Linck, J. A. (2007). Representation and skill in second language learners and proficient bilinguals. In I. Kecskes & L. Albertazzi (Eds.), *Cognitive Aspects of Bilingualism*. Berlin: Springer.
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. (2005). Models of bilingual representation and processing: looking back and to the future. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of Bilingualism* (pp. 531–553). Oxford: Oxford University Press.
Lotto, L., & de Groot, A. M. B. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning, 48*, 31–69.
Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15*, 197–237.

Markman, E. M. (1989). *Categorization and Naming in Children: Problems of Induction.* Cambridge, MA: The MIT Press.

Masson, M. E. J. (1995). A distributed model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 3–23.

McClelland, J., & Rumelhart, D. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 114*, 159–188.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227–234.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology-General, 106*, 226–254.

Perea, M., Dunabeitia, J. A., & Carreiras, M. (2008). Masked associative/semantic priming effects across languages with highly proficient bilinguals. *Journal of Memory and Language, 58*, 916–930.

Perfetti, C. A., Wlotko, E. W., & Hart, L. A. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning Memory and Cognition, 31*, 1281–1292.

Potter, M. C., So, K. F., Von Eckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23*, 23–38.

Pressley, M., Levin, J., & McDaniel, M. A. (1987). Remembering versus inferring what a word means: Mnemonic and contextual approaches. In G. McKeown & M. E. Curtis (Eds.), *The Nature of Vocabulary Acquisition* (pp. 107–127). Hillsdale, NJ: Lawrence Erlbaum Associates.

Silverberg, S., & Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language, 51*, 381–398.

Sternberg, R. J. (1987). Most vocabulary is learned from context. In G. McKeown & M. E. Curtis (Eds.), *The Nature of Vocabulary Acquisition* (pp. 89–105). Hillsdale, NJ: Lawrence Erlbaum Associates.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 12530–12535.

Tulving, E. (2001). Episodic memory and common sense: How far apart? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 356*, 1505–1515.

Tzelgov, J., & Eben-ezra, S. (1992). Components of the between-language semantic priming effect. *European Journal of Cognitive Psychology, 4*, 253–272.

van Hell, J. G., & de Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1*, 193–211.

Weinreich, U. (1953). *Languages in Contact*. New York, NY: The Linguistics Circle of New York.

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.