The Bilingual Lexicon

John N. Williams

Department of Theoretical and Applied Linguistics, University of Cambridge

A slightly revised version of this manuscript appeared in *The Oxford Handbook of the Word*, J. Taylor (Ed). OUP. 2015.

Abstract

Psycholinguistic research provides a wealth of evidence that when performing tasks in one language bilinguals and proficient second language learners cannot avoid activating orthographic, phonological, lemma, and semantic representations in their other languages. These other-language influences are evident in performance measures such as reaction time, eye movements, and brain potentials. Representations in a bilingual’s different languages continuously compete with each other for selection, suggesting that they are stored within compound systems. This is the case both for early simultaneous acquirers and adult second language learners. With regard to mapping form onto meaning, less proficient second language learners tend to rely on direct connections to L1 translations, whilst acquiring direct language-specific mappings from form to meaning requires a large amount of experience. Bilinguals rely on domain-general control mechanisms to manage the activation levels of their different languages.

Keywords: bilingualism, psycholinguistics, second language acquisition, word recognition, semantic representation, translation, executive control
Introduction

The essential questions motivating much psycholinguistic research on the bilingual lexicon go back to Weinreich’s (1953) classification of different types of bilingualism which he linked to different acquisition contexts. Drawing on Saussure’s distinction between the signifier (word) and the signified (concept) he proposed that the child exposed to two languages from birth represents their two languages in a common, fused system of signifier-signified relations. Words in the two languages access a common conceptual store. Weinreich referred to this as ‘Compound Bilingualism’. In contrast the adult learning a second language stores their languages in separate systems of signifier-signified relations leading to ‘Coordinate Bilingualism’. Alternatively, the second language is parasitic on the first, and second language words only activate meaning via their first language translations following associations between signifiers, known as ‘Subordinative Bilingualism’. Much the same distinctions have driven psycholinguistic research on the bilingual lexicon. Can bilinguals keep the representation and processing of their languages distinct, suggesting a co-ordinate organisation? Or is there interaction between their languages in moment-by-moment processing and overlap in underlying representation, suggesting a compound organisation? And how does the representation and processing of second language words change as a function of proficiency and acquisition context? Although not all of the possible combinations of acquisition context, proficiency, and modality (listening, reading, speaking) have been tested, it is already clear that the situation is not quite as Weinreich imagined.

Throughout this chapter the term “bilingual” will be used very loosely to refer to anyone with knowledge of more than one language. For example, an “advanced Dutch-English bilingual” would be someone whose first language (L1) is Dutch and who speaks English as a second (or even third or fourth) language (“L2”) to an advanced level. The research described here has been on people who, unless otherwise stated, began to learn to their second language after their first, usually in a school context, and who are at an advanced level of proficiency. Typically the participants are University students, tested either in their home country or while studying abroad.
Comprehension Processes

Recognising forms

First, consider the simple act of recognising a spoken word out of context. If you are a coordinate bilingual you should behave essentially like a monolingual in each of your languages. When listening to your L1 you will engage your L1 system, and when listening to L2 you will engage your L2 system. But this is not the case. Suppose a Dutch-English bilingual hears the phrase “click on the desk” in the context of a display containing pictures of a desk, a dustbin lid, a pineapple, and a globe. The experiment is conducted entirely in English. Their eye movements are monitored and reveal they are likely to glance at the dustbin lid before settling on the desk (Weber & Cutler, 2004). Monolingual speakers of English are much less likely to do this, so why do the Dutch? The Dutch for ‘lid’ is deksel, so it appears that on hearing the initial /de/ of “desk” the bilinguals momentarily activate the Dutch word, even though they know that they will only hear English words in the experiment (see also Spivey & Marian, 1999, for a similar effect in Russian-English bilinguals). They do not seem to be able to prevent auditory input contacting their Dutch lexicon and affecting where they direct their attention, even though they know that no Dutch words will occur. In fact, what we see here is no more than an extension of the normal processes of parallel activation of multiple lexical candidates, competition, and resolution that occur during monolingual listening (Marslen-Wilson, 1989; McClelland & Elman, 1986). Whilst all listeners have to resolve competition between similar sounding words during the recognition process, the bilingual has to deal with competitors from their other language(s) potentially delaying recognition, or at least making recognition more effortful. These results also show that input is being projected to representations in both the current and non-current languages simultaneously. A similar effect can be obtained when the heard word is in the L1 and the competitor comes from a weaker L2, although only when the competitor is a phonologically similar cognate of an L1 word (Blumenfeld & Marian, 2007).
Research on visual word recognition points to the same conclusion. If a Dutch-English bilingual is performing a lexical decision task⁠¹ in English then their decisions will be faster than expected for words like TYPE, which are identical to the Dutch word with the same meaning but differ slightly in pronunciation (Dijkstra, Grainger, & van Heuven, 1999). Facilitation is even obtained when the cognate is only similarly spelled in the two languages, such as the English-Dutch pair flood-vloed (Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010). These are true, or near, cognates, and the facilitation is assumed to reflect activation of L1 representations during a task that is conducted entirely in the L2. For ‘false friends’ like RUST (which means ‘rest’ in Dutch – same form, different meaning) lexical decision times are slower than expected, but only if the list in which they are embedded also contains words from the other language (Dijkstra, Van Jaarsveld, & Ten Brinke, 1998; Smits, Martensen, Dijkstra, & Sandra, 2006). This is presumably because the form activates different meanings, or alternatively, because different representations of orthographic word form become active in each language and then have to compete with each other for control of the required response (Dijkstra & Van Heuven, 2002). Again these effects are merely an extension of the normal recognition processes in the monolingual. A written word momentarily activates representations of orthographically similar words from the “bottom up”, and this competition is resolved as more stimulus information arrives (McClelland & Rumelhart, 1981). A consequence of this competition between alternatives in the monolingual is that the time to recognise a word is affected by how many similarly-spelled words there are in the language – the more such words, the more intense the competition for recognition (Johnson & Pugh, 1994). In the bilingual, the number of similarly spelled words in all languages known to the participant is the best predictor of recognition time (van Heuven, Dijkstra, & Grainger, 1998). This is an important result because competition between words across languages implies that functionally speaking their lexical representations are contained within the same system. Thus, recognition processes operate in essentially the same way in the monolingual and bilingual, it is just that the bilingual cannot avoid accessing representations in non-current languages.

One may object that these tasks require responses to individual words. What about the more natural situation of hearing or reading words in context? There is

⁠¹ In a lexical decision task the participant has to indicate by pressing response keys whether a stimulus is an English word or a pseudoword (e.g. BLEMP) as quickly but as accurately as they can.
evidence that in highly constraining contexts, bilinguals do have language-specific expectations for upcoming words. So a Spanish-English bilingual reads the code-switched word *carta* (‘letter’) more slowly in *He needed to put a stamp on the carta before he mailed it*, than in the more neutral *Andrea dropped a carta in the mailbox on the corner*, whereas the opposite pattern is found if the critical word is in English (Altarriba, Kroll, Sholl, & Rayner, 1996). This suggests that the linguistic context does lead to higher activation of words in the expected language when the context is itself highly constraining. On the other hand, using eye movement tracking during reading sentences in the L2, Libben & Titone (2009) found the typical relatively fast reading of cognates and inhibition of false friends in neutral sentence contexts for highly proficient French-English bilinguals. This shows that merely being embedded in an utterance of the same language does not by itself induce language-selective lexical access. In highly constraining contexts they only found cognate effects in the early stage recognition measures (first fixation time), suggesting that, as in Altarriba et al. (1996), when specific predictions can be made, they are language-specific. But in this case they reduce, and do not eliminate, the tendency to access representations in the non-current language. And using a priming methodology on German-English bilinguals, Elston-Guttler et al. (2005) found that the word *gift* in “The woman gave her friend an expensive gift” primed POISON (*gift* means poison in German) showing that the word had activated its German meaning even though this was irrelevant in the context. These latter two studies therefore suggest that even constraining sentence contexts do not necessarily eliminate cognate effects in bilinguals. Elston-Guttler et al.’s participants were advanced learners of English tested in Germany, and Libben & Titone’s participants were highly proficient French-English bilinguals tested in Montreal, so in both cases the L1 was likely to be used outside of the laboratory, and hence not dormant. Interestingly, in Elston-Guttler et al., if the participants first watched a 20-minute movie in their L2 English just before doing the task then the cognate priming effect disappeared, suggesting that language context can modulate interference from the L1 if it is sufficiently well-established. Thus, linguistic context clearly can reduce cognate effects on recognition, just as one would expect from the general interplay of bottom-up and top-down factors that characterizes language processing in general. What is perhaps surprising, though, is that a language context *per se* has little effect unless it is strongly instantiated. Further studies manipulating language context and country of testing are clearly needed.
The above studies show that word forms in current and non-current languages compete with each other during the recognition process. This suggests a ‘compound’ form of representation, since for words to compete with each other they must be stored within a common system. Models of bilingual word recognition make this assumption explicit: the Bilingual Interactive Activation, BIA, model (Dijkstra & Van Heuven, 2002) for visual word recognition, and the Bilingual Model of Lexical Access, BIMOLA, for spoken word recognition (Grosjean, 1997). For languages that share certain graphemes and phonemes it seems intuitively obvious that the same representations should be utilised, but the deeper claim is that the abstract representations of word forms are also contained in a common system. This compound form of representation appears to apply to late learners, at all levels of proficiency.

*From form to meaning. A common conceptual code?*

Following Weinreich (1953) one may wonder whether common or separate meanings are accessed by L1 and L2 words. But this question, simple as it may seem, obscures deep questions about what is meant by ‘meaning’ and what is meant by ‘common’. With regard to meaning, do we make a distinction between conceptual structure and lexical meanings – a two-level semantics (Bierwisch & Schreuder, 1992; Pavlenko, 1999), or is there simply a one-to-one mapping between concepts and words (Levelt, Roelofs, & Meyer, 1999)? Clearly it would be naïve to assume that words in different languages can have exactly the same meanings, or that all meanings are expressible in all languages. For example, Japanese has one word, *ashi*, that refers to the leg and the foot. Clearly this does not mean that Japanese people cannot conceive of legs and feet as separable concepts (Vigliocco & Vinson, 2007). The assumption in bilingual lexicon research is that whereas there exists a common, language-independent conceptual substrate that is accessed by language (and which motivates utterances), there must be language-specific mappings from concepts to words. This does not necessarily mean that lexicalisation patterns cannot affect conceptualisation (Vigliocco & Filipovic, 2004). For example, Pavlenko (1999) argues for a distinction between lexical-semantic representations and concepts, with the former being potentially rapidly acquired, and permitting language-specific conceptual structures to be formed through experience over time. In a different vein, Van Hell & De Groot
(1998) propose a distributed feature model, inspired by connectionist approaches to word meaning, in which word forms from different languages map onto a common conceptual code, and induce their own language-specific patterns of activation. As a result, words from different languages may share varying numbers of conceptual features. For example, comparisons between word associations in L1 and L2 (Van Hell & De Groot, 1998) suggest that there is more overlap between the conceptual codes activated in L1 and L2 by concrete words and cognates than abstract words. In this way a common code can be combined with language-specific patterns of activation over that code.

Evidence for a common conceptual substrate across languages comes from cross-language semantic priming. Within one language the time to recognise a word, say CHAIR, is facilitated if it is preceded by a semantically related ‘prime’ word like TABLE (Meyer & Schvaneveldt, 1971). The same effect can be obtained reliably across languages, at least when the prime is in the L1 and the target in the L2. For example, for an English-Italian bilingual TABLE will prime SEDIA (‘chair’) (see Altarriba & Basnight-Brown, 2007, for a review). In order to provide convincing evidence of shared underlying meaning such effects have to be automatic, e.g. the result of an automatic spreading activation around the conceptual system (Collins & Loftus, 1975), and not the result of conscious expectancy strategies (Neely, 1977). Bilingual researchers have used a range of techniques to ensure automaticity: tasks that distract attention from the primes (Fox, 1996); brief, but still visible primes (Keatley, Spinks, & de Gelder, 1994; Tzelgov & Ebenezra, 1992); or primes presented so briefly that they are subliminal (Basnight-Brown & Altarriba, 2007; Perea, Dunabeitia, & Carreiras, 2008; Williams, 1994). Cross-language semantic priming effects have been obtained in all cases, suggesting that L1 and L2 words do share components of meaning. Once again, even for late learners and learners at various levels of proficiency we find evidence for compound organization, in the sense that a common conceptual system is shared between languages, although without denying the possibility of language-specific conceptual structures forming within that system.

Accessing meaning from form
Even if we accept that words in all languages ultimately map onto a common conceptual code, there remains the issue of how it is accessed by L2 forms and how this may depend on level of proficiency. One highly influential hypothesis is that at low levels of proficiency L2 words access meaning via their L1 translations, an idea instantiated in the “Revised Hierarchical Model” of bilingual memory, or RHM (Kroll & Stewart, 1994; Kroll, Van Hell, Tokowicz, & Green, 2010). To begin with, L2 words have much weaker mappings onto the conceptual system than their L1 translations. So learners use direct lexical connections between L2 words and their L1 translation equivalents enabling them to exploit the strong mapping from L1 words to meaning. This is equivalent to the notion of subordinative bilingualism in the sense that the L2 is parasitic on the L1. With increasing proficiency learners develop direct mappings from L2 forms to meanings.

There are four main lines of evidence that support the RHM. First, comparisons of picture naming and word translation. At low levels of proficiency translation from L1 to L2 is faster than naming pictures in L2, whereas at higher proficiency there is no difference (Chen & Leung, 1989; Potter, So, Voneckardt, & Feldman, 1984). Naming pictures in L2 requires use of direct form-meaning connections, which is problematic at low proficiency, but translation can be achieved by exploiting L1-to-L2 lexical-level connections, and hence is relatively fast. At higher proficiency, both tasks engage the conceptual level and so take about the same time. Second, semantic blocking effects in translation. When naming pictures one after another, response times slow down when items are presented in blocks organised by semantic category, reflecting competition in the concept-form mappings. The same effect occurs when translating words from L1 to L2, suggesting conceptual mediation, but not when translating from L2 to L1, suggesting that the conceptual level is bypassed and lexical-level translation connections are used instead (Kroll & Stewart, 1994). Third, form and meaning interference in translation recognition. In this task people are given pairs of words and asked to indicate if they are correct translations. For example, an English-Spanish bilingual might be given pairs like MAN – HAMBRE or MAN – MUJER. The correct response in both cases is ‘no’, but responses might be relatively slow because HAMBRE is orthographically similar to ‘hombre’ (‘man’), and MUJER means ‘woman’. A slow-down in the first case reflects form-level interference and is greater at low than high levels of proficiency, whereas in the second case the interference is meaning-based and is greater at high proficiency.
(Talamas, Kroll, & Dufour, 1999). This pattern of results is consistent with a shift from lexical to conceptually-mediated translation, as predicted by the RHM. Fourth, semantic and translation priming effects. Whereas there is good evidence for automatic semantic priming from L1 to L2, effects tend to be much weaker or entirely absent in the L2 to L1 direction (Basnight-Brown & Altarriba, 2007; Fox, 1996; Keatley et al., 1994; Tzelgov & Ebenezra, 1992, Silverberg & Samuel, 2004). Only in highly proficient bilinguals or early acquirers of high proficiency are the semantic priming effects the same size regardless of direction (Perea et al., 2008; Silverberg & Samuel, 2004). This is consistent with L2 form-meaning mappings being initially relatively weak, as one might expect. In contrast, many of the studies that failed to find L2-L1 semantic priming did find L2-L1 translation priming; e.g., from CHIEN to DOG for an English-French bilingual (Basnight-Brown & Altarriba, 2007; Fox, 1996; Keatley et al., 1994) and L2-L1 translation priming has even been obtained for newly learned words (Altarriba & Knickerbocker, 2011; Williams & Cheung, 2011). The weakness of L2-L1 semantic priming and relative robustness of L2-L1 translation priming at anything but very high levels of proficiency is consistent with the RHM.

Although many lines of evidence lend support to the RHM, it is probably an over-simplification to say that all L2 words initially activate meaning only via L1 translations. If translation is achieved purely lexically then translation times should not be affected by semantic variables. But De Groot & Poot (1997) found that Dutch-English bilinguals translate concrete words more quickly than abstract words in both translation directions and at all levels of proficiency. This points to the involvement of meaning in the translation process. And in translation recognition tasks, meaning-based interference has been found even in low proficiency bilinguals (Altarriba & Mathis, 1997; Sunderman & Kroll, 2006). Given the emerging picture from psycholinguistics for high levels of interaction and cross-talk between different codes it seems unlikely that bilinguals of whatever proficiency would ever rely on a single pathway for accessing meaning from L2 words. It seems more likely that activation traverses multiple pathways, with different degrees of involvement of different pathways according to proficiency and type of word.

In fact, reflexes of translation effects can emerge even at quite high levels of proficiency. Silverberg & Samuel (1994) showed that for early Spanish-English bilinguals BULL facilitated lexical decisions on the Spanish target TORNILLO (‘screw’). They argue that this is because BULL activates its meaning directly, which
then activates the form of the Spanish translation toro, which then facilitates recognition of the form-similar TORNILLO. Thierry & Wu (2007) gave proficient late-learner Chinese-English bilinguals a semantic relatedness judgement test including English word pairs like Post-Mail and Train-Ham. Some of the pairs would contain a character repetition when translated into Chinese. For example the pair Train-Ham translates as Huo Che – Huo Tui (火车 -火腿). Different ERP responses were obtained for items with and without character repetition, suggesting that at an unconscious level the participants were being influenced by the form similarity between the words in their native language. Even though the task was all in English they were implicitly translating the words into Chinese. Wu & Thierry (2010) provide evidence that, specifically, they are being influenced by the sound of the Chinese translations, and not their spelling. Thus, there are reverberating effects around the lexical system at all levels reflecting interactions between translation equivalents, although whether these necessarily reflect the kinds of translation connections posited by the RHM is not clear.

What meaning is accessed by L2 words?

The research reviewed above intentionally does not address the complexities of form-meaning mappings across languages. The notion of ‘translation equivalence’ is treated simplistically as items sharing a common meaning. Clearly this is not the case for many words. For example, a homonym in one language (e.g., bank in English) will most likely be translated by different words in another (e.g., in Italian, banca for financial bank and riva for river bank). One-to-many mappings in meanings across languages slows down both processing and acquisition due to competition between alternative translations (Degani & Tokowicz, 2010).

A more subtle problem occurs when L2 words correspond to polysemous meanings of an L1 word. For example, the Chinese word jie covers the meanings of both English lend and borrow, and wenti covers the meanings of problem and question. Jiang (2002) showed that Chinese-English bilinguals rate word pairs like problem-question as more highly semantically related (relative to unrelated control pairs) than do English monolinguals, and their response times are faster. Jiang argues
that this is because they essentially utilise the L1 meaning of the L2 translation, and are only prevented from rating the words as having an identical meaning by explicit knowledge of their non-equivalence.

It does appear that L1 polysemy continues to colour the interpretation of L2 words even at high levels of proficiency. In Elston-Guttler & Williams (2008), advanced German-English bilinguals were asked to indicate whether the last word of sentences formed a plausible completion. For a sentence such as “His shoes were uncomfortable due to a bubble” the time to respond ‘Implausible’ was slower than for native speakers of English (relative to the plausible sentences). The reason is that the translation equivalent of bubble in German is Blase, which also means ‘blister’. Blase would not be regarded as a homonym, but rather as a polysemous word that has a broader meaning than the English translation equivalent. Once again it appears that the L1 meaning influences L2 processing. However, note that in this case the majority of responses were correct. The learners did indicate that the completion was implausible, suggesting that they knew the correct English meaning of the word. Either this is a reflection of explicit knowledge of the correct meanings, or a reflection of residual connections in the system from earlier stages of development, whether lexical level translation connections or connections from L2 form to L1 meaning. In any case, the shadow of L1 meaning appears to be remarkably difficult to shake off.

Given the existence of homonomy and polysemy within all languages, it is more realistic to think in terms of acquisition of senses, rather than a unitary meaning. Whilst learners may indeed assume that words that are learned as translation equivalents share a common core meaning, they may not assume that all of the senses of the word are shared. Kellerman (1979) found that learners’ willingness to transfer a particular usage of the L1 Dutch word breken (to break) to the L2 was related to how close it was to the core meaning of the word, as established through multidimensional scaling. For example, ‘break a cup’ (in Dutch) was rated as being readily translatable into English, whereas ‘break a fall’ was not; clearly the former is more representative of the core meaning of breken. Thus, the various senses of an L2 word may need to be acquired through experience, and not simply copied over from an assumed translation equivalent. Finkbeiner & Nicol (2004) argue that if L2 words are associated with fewer senses than L1 words then this can explain the asymmetry in priming effects noted earlier, i.e. that priming effects tend to be weaker from the L2 to the L1 than vice versa. Indeed, even priming between translations can show this asymmetry when
primes are masked and the target task is lexical decision (Finkbeiner et al., 2004; Jiang & Forster, 2001). If one considers the proportion of the target meaning that overlaps with the prime, an L2 target with few senses will overlap completely with an L1 prime with many senses. But an L1 target with many senses will only partially overlap with an L2 prime with few senses. Thus, acquisition of word meaning over time can be conceived as the accumulation of senses of the word.

Even within a single word sense, one might wonder which specific components of meaning are shared between translations, and whether L2 words simply inherit the entire meaning of the L1 sense with which it is associated. Williams (1994) provides evidence that automatic (in this case subliminal) L1 to L2 priming is most reliable for pairs that are semantically similar, e.g. FENCE – HAIE (‘hedge’ in French), as opposed to merely associated, e.g. NEEDLE – FIL (‘thread’ in French). De Groot & Nas (1991) also failed to find subliminal cross-language priming for associates. This may suggest that L2 words do not necessarily inherit the conceptual associates of L2 translations, but map onto a more restricted lexical meaning (Pavlenko, 1999). This point will be further elaborated below.

How can direct form-meaning connections be acquired?

If acquiring direct form-meaning connections in the L2 is problematic, what can be done to help? There is some evidence that even at the early stages of vocabulary learning differences in instructional method can have an effect. Comesana et al. (2009) showed that meaning interference effects in translation recognition, indicative of direct meaning access, could be obtained after only one session of instruction on novel words, but only if the words were learned with pictures. Perhaps, then, if the instructional technique can avoid reliance on translations direct mappings to meaning can be acquired more easily. On a larger scale, Linck, Kroll, & Sunderman (2009) found meaning interference effects in translation recognition only for immersion learners, and not for classroom learners matched for proficiency. Silverberg & Samuel (2004) only found automatic L2-L1 semantic priming for early learners, who had also learned their L2 through immersion. This research points to the role of extensive language experience in real contexts of use for developing strong form-meaning mappings in the L2.
In a laboratory-based study of word learning, Williams & Cheung (2011) examined L3 to L1 semantic priming effects from newly learned words in Chinese-English bilinguals learning French as L3. Recall that previous research has generally shown L2 to L1 semantic priming effects to be weak. In this study, L3 (French) words were first learned with their L2 (English) translations, for example écureuil was learned as the translation of squirrel. The participants then performed a semantic priming test using the novel L3 (French) words as primes and L1 (Chinese) words as targets. Robust priming of translation targets was obtained, e.g., from écureuil to 松鼠 (squirrel in Chinese). Since the French words were never learned with their Chinese translations this priming is not due to mere repetition of the learning experiences, but presumably reflects a deeper semantic connection. This was confirmed by priming from the French words to Chinese targets that were semantically similar, e.g., between écureuil and 松鼠 (‘animal’). However, no priming was obtained to targets that were associated but not semantically similar, e.g. between écureuil and 果仁 (‘nut’). This suggests some restriction on the semantic information that is accessible from words when learned in translation pairs and is reminiscent of the previous priming results of Williams (1994), which also showed a dissociation between priming for semantically similar and associated pairs. In Williams & Cheung (2011) the only learning task that resulted in priming of associates was one in which the novel French words were learned in the context of pictures that actually contained the associate. For example, some of the pictures depicted a squirrel holding a nut. In the subsequent test, écureuil primed 果仁 (‘nut’). These results suggest an essentially episodic view of word learning (see also Perfetti, Wlotko, & Hart, 2005). Novel words may readily map onto the lexical meaning of their translation equivalents, but they may form associations with other concepts only through experience.

Pavlenko (1999) makes a distinction between lexical meanings and the conceptual structures to which they ultimately relate and which provide words with, amongst other things, connotative meaning. She quotes from Eva Hoffman's (1989) book Lost in Translation: A Life in a New Language:

The words I learn now don't stand for things in the same unquestioned way they did in my native tongue. 'River' in Polish was a vital sound, energized with the essence of riverhood, of my rivers, of my being immersed in rivers. 'River' in English is cold – a word without an aura. It has no accumulated
associations for me, and it does not give off the radiating haze of connotation. It does not evoke.
(Hoffmann 1989: 106)

Thus, while learners may rapidly learn a mapping between an L2 form and a lexical meaning inherited from an L1 translation (Jiang, 2002), only through episodes of actual usage will they acquire the associations between this word meaning and deeper conceptual structures. More needs to be done to explore the impact of learning experiences on the underlying representation of word meaning and to tease apart different aspects of meaning in a more systematic way than is usual in this research.

Production

The theme running through all of the above comprehension studies is that bilinguals show detectable processing effects due to activation of lexical and semantic representations in the non-current language. Bilinguals do not appear to be able to entirely “switch off” their other languages, even in situations where only one language appears to be relevant. Perhaps this is not too surprising since it may be difficult to apply a language filter to determine which pathways an incoming stimulus will traverse. But when it comes to production, apart from code-switching or occasional other-language intrusions (Poulisse & Bongaerts, 1994), bilinguals do generally appear to be able to speak in the intended language, implying that they are able to switch off their other languages. However, psycholinguistic research has revealed that here too activation of representations in the non-concurrent language has a detectable effect.

In the case of visual word recognition we saw that processing of a word in one language is affected by cognates in other languages. Similar cognate facilitation effects occur in production tasks, such as picture naming. If an advanced Catalan-Spanish bilingual is asked to name pictures in Spanish they are faster to do so for cognates like *gato* (*gat* in Catalan, ‘cat’) than for non-cognates like *mesa* (*taula* in Catalan, ‘table’). The effect can also be obtained when pictures are to be named in the L1, although the effect is smaller (Costa, Caramazza, & Sebastian-Galles, 2000). Similar effects on L2 picture naming have been obtained for Japanese-English
bilinguals, where the scripts are very different (Hoshino & Kroll, 2008). These results suggest that when a bilingual wants to express a concept in the current language they cannot avoid activating the phonological form of the word in non-current languages. For cognates the overlap in phonological representations produces faster responses. Colomé (2001) showed that effects of phonological activation in the non-current language can be detected for non-cognates as well. Highly advanced Catalan-Spanish bilinguals were shown a letter, e.g., m, followed by a picture, e.g., a table. They had to indicate whether the sound of the letter was contained in the name of the picture in Catalan, but without saying the name of the picture aloud. In this case the correct answer is “no” because the name of the picture is taula. However, the Spanish for table, mesa, does contain the target sound, and this makes people slower than for trials where the target sound is in neither the Catalan or Spanish names. Participants seemed to automatically activate the sound of the Spanish translation of the word even though this was completely irrelevant to the task. There is therefore evidence that even in production tasks bilinguals cannot switch off non-current languages.

What about mapping from concepts to words (the reverse of form-meaning mapping in comprehension)? It is known from research on monolinguals that when naming an object there is competition between lexical alternatives (Levelt et al., 1999). For example, Wheeldon & Monsell (1994) showed that if a person has just said “shark” in response to the definition ‘Man-eating fish’, a few trials later their time to name a picture of a whale will be slower than if they had not recently produced a related word. This effect is not due to problems recognising the object but to finding its name, i.e., the process of lexicalization. It is assumed that whenever one entertains the concept of a shark, words with similar meanings are also activated to some extent and compete with the intended word for selection (a similar process is more obviously at work in word substitution errors of the kind “Put it in the oven at a low speed”). Prior production of a potential competitor temporarily increases its activation level and it interferes with production of the intended word more than usual. This effect shows that concepts are mapped onto words in a competitive process (see also Schriefers, Meyer, & Levelt, 1990, for evidence from a picture-word interference paradigm). From a bilingual perspective, one can then ask whether this competition extends to words in the non-current language. There is research to suggest that it does. Using the same paradigm as Wheeldon & Monsell (1994), Lee & Williams (2001) showed that if an English-French bilingual utters “shark” in response to a definition, a
few trials later they will suffer interference in naming a picture of a whale in French as “baleine” (see also Hermans, Bongaerts, De Bot, & Schreuder, 1998, for evidence from a picture-word interference paradigm). This result suggests that words in the non-intended language compete for selection in production with words in the intended language. It is consistent with the research described above showing activation of non-current language words during production, but also provides strong evidence that the relevant representations are contained within the same system. It should be noted though that Costa et al. (1999) have claimed that whilst there may be activation of non-current language words these do not actually compete for selection with words in the intended language. They showed that when a Catalan-Spanish bilingual names a picture of a table in L1 as “taula” their responses are facilitated if the word *mesa* (which means ‘table’ in Spanish) is printed on the picture, compared to an unrelated control word. If there really were competition between words in different languages for selection then a large interference effect from *mesa* would have been expected. Clearly this is a very different, and even less natural, paradigm from that used by Lee & Williams (2001), and possibly behaviour was affected by strong L2 to L1 translation connections, as proposed by the RHM. Or else the ability to avoid competition from other-language words might reflect a change in the language control mechanism at higher levels of proficiency (Costa & Santesteban, 2004).

In order to explain competition effects between alternatives during lexical selection it is usual to appeal to a “lemma” level of representation that is intermediate between concepts and word forms. This controls the mapping between lexical concepts and forms and is the locus for attaching syntactic information about the word (Levelt et al., 1999). For example, lemmas for verbs will be linked to representations indicating whether they are transitive or intransitive, or whether they participate in ditransitive and/or prepositional dative constructions. Nouns in certain languages will be linked to representations indicating their grammatical gender. There is evidence that in bilinguals there is an interaction between the syntactic information attached to lemmas in the two languages. For example, Greek-German bilinguals are faster to translate noun phrases (containing a gender marked adjective followed by a noun) from L1 Greek into L2 German if they have the same grammatical gender in both languages (Salamoura & Williams, 2008). Similarly, German-Dutch bilinguals are faster to name pictures in their L2 Dutch if the grammatical gender of the noun is the same in the two languages (Lemhöfer, Spalek, & Schriefers, 2008). In the case of
argument structure alternations, Salamoura & Williams (2007) showed that if a Greek-English bilingual uses, say, a double object dative structure in L1 Greek, then they are relatively likely to reuse that structure subsequently when producing an utterance in L2 English, replicating the usual within-language syntactic priming effect (see also Schoonbaert, Hartsuiker, & Pickering, 2007). Cross-language gender and syntactic priming effects can be regarded as reflecting the sharing of syntactic representations associated with lemmas in different languages, and again point to the relevant representations being contained in essentially the same system.

Despite there being activation of lexical representations in the non-current language even during production, people do nevertheless generally speak in the intended language. Obviously there is some additional control process at work to ensure that the intended language is used. One influential proposal about how this is achieved is through inhibition of the non-current language (Abutalebi & Green, 2007; Green, 1998). Lexical representations (specifically lemmas in this model) carry language tags that allow words in specific languages to be inhibited according to task demands. Support for this idea comes from the observation that when switching between languages there is a greater cost associated with switching from a weaker L2 to the stronger L1 than vice versa (Lee & Williams, 2001; Macnamara, Krautham, & Bolgar, 1968; Meuter & Allport, 1999). For example, an English-French bilingual might be asked to name two pictures in succession, with the language to be used for each picture indicated by a cue. If the first picture is named in L2 French, the time to name the second picture in L1 English will be slower than if the first picture had also been named in L1 English. This switching cost is much greater than for the comparison between English-French and French-French trial sequences (Lee & Williams, 2001). The reason for this is that in order to speak a relatively weak L2 the strong L1 has to be massively inhibited. If the L1 is then required it takes relatively more time to disinhibit it. But in order to speak the L1 a weak L2 requires hardly any inhibition, and hence, paradoxically, is relatively more available when required. The executive functions that are at work here are not specific to language, but engaged in all task-switching situations (Macnamara et al., 1968; Meuter & Allport, 1999). The switching cost asymmetry is not found in balanced bilinguals, which according to the inhibitory control account is because both languages require equal amounts of inhibition. Alternatively it might reflect a greater ability to simply ignore activated
lexical representations in the non-current language, implying a change in the control mechanism with increasing proficiency (Costa & Santesteban, 2004)

Conclusion

Bilinguals are unable to simply switch off their non-current languages. Experiment after experiment shows that L1 and L2 lexical representations are simultaneously active during processing in both comprehension and production, leading to various “cross-talk” effects between the languages. In recognition and production non-current language forms become active, producing cognate facilitation and false friend interference in recognition tasks. In the mapping between form and meaning in comprehension the meanings of other-language translations colour the semantic activity produced by L2 words. In the mapping between meaning and form in production there is activation of other language lemmas. Thus, bilinguals appear to have surprisingly little control over their language systems, particularly in comprehension. To be sure, a control mechanism does exist, and is most evident in production tasks. Exercising this control system to manage their languages may enhance executive functions in bilinguals (Bialystok, 2009). The fact that language-independent brain areas for task management have to be recruited implies that language-internal mechanisms for managing the activation of the competing language systems are not sufficient.

But co-activation of a bilingual’s language systems does not necessarily tell us anything about their underlying organisation. Even if the L1 and L2 lexical representations were contained in distinct systems, as in co-ordinate bilingualism, they could still become simultaneously active in response to linguistic stimuli or communicative intentions. What evidence is there that lexical representations in different languages are actually contained within the same system? This is where direct evidence for competition between representations in different languages is critical because it implies that representations are in some sense contained within the same functional system, else why would they have to compete with each other? Here we can point to evidence that in the bilingual the orthographic neighbourhood for visual word recognition has to be defined over both L1 and L2, and to cross-language competition effects in word production. Cross-language gender and syntactic priming effects suggest that lemma representations for words in different languages are
connected to shared representations of their syntactic properties. Such findings point to a compound representation of lexical knowledge. Whereas Weinreich (1953) thought that this was characteristic of early, simultaneous acquirers, instead we find compound organisation even for later learners of second languages.

The compound view is also consistent with neuroimaging work on lexico-semantic processing in bilinguals that shows activation in similar brain areas (specifically left frontal and temporo-parietal cortex) in L1 and L2 across a range of proficiency levels and lexical tasks (for reviews see Abutalebi, 2008; Indefrey, 2006). This overlap has been observed even for languages as dissimilar as Chinese and English (Chee, Tan, & Thiel, 1999). The only differences between L1 and L2 brain activity that do occur are at lower L2 proficiency, where there is additional brain activity in pre-frontal areas, regions known to be associated with domain-general executive functions. This is assumed to reflect the additional control demands required when processes are of low automaticity, and as discussed above, when the L1 has to be inhibited in order to speak a weaker L2 (Abutalebi, 2008).

A somewhat different picture emerges when we consider the mapping between form and meaning in comprehension. Here there is good evidence for a subordinative relationship between L2 and L1, at least in the earlier stages of acquisition; L2 words access meaning via a lexical-level translation connection to the L1, enabling them to exploit the strong mapping between L1 form and meaning. However, it is also clear that all-or-none, system-wide statements of this kind are an oversimplification. Mediated access to meaning is merely a tendency; there is also evidence for direct access even from newly learned words, and a greater semantic involvement in translation for some words rather than others. Simple box-and-arrow models such as the RHM have been useful in allowing researchers to formulate hypotheses about access routes to meaning, but it is now time to move beyond gross distinctions between lexical form and meaning, and address at a finer level of grain the actual content of the semantic and conceptual representations that are accessed by L2 words, how this varies at different levels of proficiency, and perhaps most importantly of all, how it is affected by different learning tasks and environments. After all, the way that lexical knowledge is represented in the mind of the bilingual is a result of learning experiences, and it is through understanding the connection between learning, representation, and processing that this field can have more impact upon teaching practice and language policy.
References


